

Sat Nov 6 18:59:24 2004

us-10-054-873-4.rag

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 2, 2004, 19:47:31; Search time 82.9852 Seconds
(without alignments)
371.762 Million cells updates/sec

Title: US-10-054-873-4
Perfect score: 463
Sequence: 1 FVNHGCHSLVYALYVCG.....IVEQCCTSLCSLVLENYCN 86

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: A_Geneseq_23Sep04:*

- 1: Geneseq21980s:*
- 2: Geneseq21990s:*
- 3: Geneseq22000s:*
- 4: Geneseq22001s:*
- 5: Geneseq22002s:*
- 6: Geneseq22003as:*
- 7: Geneseq22003bs:*
- 8: Geneseq22004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	463	100.0	86	1 AAP40829	Sequence
2	463	100.0	86	2 AAP40829	Sequence
3	463	100.0	86	3 AAP40829	Sequence
4	463	100.0	86	4 AAP40829	Sequence
5	463	100.0	86	5 AAP40829	Sequence
6	463	100.0	86	6 AAP40829	Sequence
7	463	100.0	86	7 AAP40829	Sequence
8	463	100.0	86	8 AAP40829	Sequence
9	463	100.0	86	9 AAP40829	Sequence
10	463	100.0	86	10 AAP40829	Sequence
11	463	100.0	86	11 AAP40829	Sequence
12	463	100.0	86	12 AAP40829	Sequence
13	463	100.0	86	13 AAP40829	Sequence
14	463	100.0	86	14 AAP40829	Sequence
15	463	100.0	86	15 AAP40829	Sequence
16	463	100.0	86	16 AAP40829	Sequence
17	463	100.0	86	17 AAP40829	Sequence
18	463	100.0	86	18 AAP40829	Sequence
19	463	100.0	86	19 AAP40829	Sequence
20	463	100.0	86	20 AAP40829	Sequence
21	463	100.0	86	21 AAP40829	Sequence
22	463	100.0	86	22 AAP40829	Sequence
23	463	100.0	86	23 AAP40829	Sequence
24	463	100.0	86	24 AAP40829	Sequence
25	463	100.0	86	25 AAP40829	Sequence

26	463	100.0	110	1 AAP10053	Sequence
27	463	100.0	110	2 AAP40309	Sequence
28	463	100.0	110	3 AAP06608	Human pre
29	463	100.0	110	4 ADF77582	Human ins
30	463	100.0	110	5 AAY44367	Human pro
31	463	100.0	110	6 AAY70366	Human ins
32	463	100.0	110	7 AAB26765	Human pre
33	463	100.0	110	8 AAB06144	Human ins
34	463	100.0	110	9 AAE10337	Human pre
35	463	100.0	110	10 AAB35424	Secretory
36	463	100.0	110	11 AAG65677	Human pro
37	463	100.0	110	12 ABG60634	Human ins
38	463	100.0	110	13 ABG31590	Human pre
39	463	100.0	110	14 ABR55862	Humanised
40	463	100.0	110	15 ADA09218	Human pre
41	463	100.0	110	16 ADC51569	Human pro
42	463	100.0	110	17 ADE56710	Human pro
43	463	100.0	110	18 ADD46938	Human pro
44	463	100.0	110	19 ADE57650	Human pro
45	463	100.0	110	20 ADJ25930	Human pre

ALIGNMENTS

RESULT 1	AAP40829	AAP40829 standard; protein; 86 AA.
XX	XX	
AC	XX	AAP40829;
XX	XX	
DT	XX	09-SEP-2004 (revised)
DT	XX	25-MAR-2003 (revised)
DT	XX	03-AUG-1992 (first entry)
XX	XX	
DE	XX	Sequence of human insulin precursor.
XX	XX	
KM	XX	Insulin precursor; connecting peptide; diabetes; hormone.
XX	XX	
OS	XX	Homo sapiens.
OS	XX	Unidentified.
XX	XX	
EH	XX	Key
FT	XX	Region
FT	XX	Modified-site
FT	XX	1
FT	XX	/label= chain B
FT	XX	/label= F-NH2-R
FT	XX	/label= "H or a chemically or enzymatically cleavable AA residue or peptide residue"
FT	XX	7. .72
FT	XX	19. .85
FT	XX	31. .65
FT	XX	/label= connecting peptide
FT	XX	66. .86
FT	XX	/label= chain A
FT	XX	71. .76
FT	XX	/label= N-OH
FT	XX	
PN	XX	US4430266-A.
XX	XX	
XX	XX	07-FEB-1984.
PD	XX	
PD	XX	
PF	XX	16-FEB-1982; 82US-00349397.
PF	XX	
PR	XX	27-MAR-1980; 80US-00134385.
PR	XX	28-NOV-1980; 80US-00210696.
XX	XX	
XX	XX	(ELI) LILLY & CO ELI.
PI	XX	Frank BH;
XX	XX	
XX	XX	WPI; 1984-049032/08.

XX Insulin precursor prodn. from linear S-sulphonate and mercaptan - in
PT single step without separate oxidn.
XX
PS Claim 17; Col 4; 8pp; English.
XX
CC The inventors claim a method for the prepn. of an insulin precursor in
CC which the A-chain and B-chain are joined through a connecting peptide.
CC The connecting peptide joins the A-chain at the amino group of A-1 to the
CC B-chain at the carboxyl group of B-30. The method is pref. for the prepn.
CC of human insulin precursor (see AAP40829). The SOS of the connecting
CC peptides of a number of species are given (see AAP40828, AAP40830-35).
CC (Updated on 25-MAR-2003 to correct PA field.)
CC
CC Revised record issued on 03-SEP-2004 : Correction to Feature Table Key
XX
SQ Sequence 86 AA;
XX
Query Match 100.0%; Score 463; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 1,4e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FVNQHLCGSHLVEAFLYGSGRFPYKPTREARLDLVGVGGLGSPGAGSLQPLALEG 60
Db 1 FVNQHLCGSHLVEAFLYGSGRFPYKPTREARLDLVGVGGLGSPGAGSLQPLALEG 60
QY 61 SLOKRGIVQCCCTSICTSLYLENYCN 86
Db 61 SLOKRGIVQCCCTSICTSLYLENYCN 86
DE 22-AUG-1996 (first entry)
AC AAR84061;
XX AAR84061 standard; protein; 86 AA.
XX
DE 22-AUG-1996 (first entry)
XX
DE Human insulin.
XX
KM Insulin; transformation; gene expression; fungi; fungal cell; hormone;
KM A-chain; C-chain; glycosylation.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..261
FT /*tag= a
FT /product= "insulin."
XX
XX EP704527-A2.
XX
XX 03-APR-1996.
XX
XX 03-AUG-1995; 95EP-00112210.
XX
XX 05-AUG-1994; 94HR-00000432.
XX
XX (PLIV) PLIVA PHARM & CHEM FAB.
XX
XX Nestric S, Punt PJ, Vallingner R, Van Den Hondel CAMU;
XX
XX WPI; 1996-129917/18.
XX
XX N-PSDB; AAT17830, AAT17831.
XX
XX DNA encoding human insulin precursors - which comprise B- and A-chains
XX linked via amino acid chain contg. 1 or more glycosylation sites, for
XX prepn. of insulin in fungal cells.
XX
XX Disclosure; Fig 1; 32pp; English.
XX
XX DNA sequences encoding insulin precursors of formula B-Pg-A, where B and

CC A represent B- and A-chains of insulin respectively, and Pg represents a
CC modified C-peptide or any number of amino acids comprising at least one
CC glycosylation consensus site, can be inserted into expression vectors
CC which in turn can be used to transform fungal host cells. The fungal
CC cells are then cultured and the insulin expressed in such cells can be
CC harvested
XX
SQ Sequence 86 AA;
XX
Query Match 100.0%; Score 463; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 1,4e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FVNQHLCGSHLVEAFLYGSGRFPYKPTREARLDLVGVGGLGSPGAGSLQPLALEG 60
Db 1 FVNQHLCGSHLVEAFLYGSGRFPYKPTREARLDLVGVGGLGSPGAGSLQPLALEG 60
QY 61 SLOKRGIVQCCCTSICTSLYLENYCN 86
Db 61 SLOKRGIVQCCCTSICTSLYLENYCN 86
DE 19-JAN-2000 (first entry)
AC AAY42858;
XX AAY42858 standard; protein; 86 AA.
XX
DE Human insulin precursor, SEQ ID 5.
XX
KM Insulin; precursor; growth hormone; chaperone; intramolecular; folding;
KM conformation; chimeric protein; cleavable; recombinant; production;
KM yield.
XX
XX Homo sapiens.
XX
XX W09950302-A1.
XX
XX 07-OCT-1999.
XX
XX 31-MAR-1998; 98WO-CN000052.
XX
XX 31-MAR-1998; 98WO-CN000052.
XX
XX (TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.
XX
XX Gan Z;
XX
XX WPI; 1999-610839/52.
XX
XX New chimeric proteins containing human growth hormone fragment, used
XX particularly for the production of human insulin.
XX
XX Claim 10; Page 29; 46pp; English.
XX
XX This sequence represents a human insulin precursor comprising insulin A
XX and B chains separated by a 34 residue peptide sequence. This insulin
XX precursor can be a component of chimeric proteins which additionally
XX contains an N-terminal fragment of human growth hormone (hGH) and a
XX cleavable peptide linker (AAY42857). The hGH portion of the chimeric
XX protein acts as an intramolecular chaperone (IMC) for the insulin
XX precursor, enabling it to fold correctly. The cleavable peptide linker
XX has a C-terminal Arg residue which enables the hGH portion of the
XX chimeric protein to be removed after folding has taken place. Production
XX of recombinant human insulin via an hGH-proinsulin chimeric protein can
XX provide human insulin with correctly linked cysteine bridges with fewer
XX necessary procedural steps, and hence resulting in a higher yield of
XX human insulin. The IMC sequences not only protect insulin sequences from
XX intracellular degradation by a microorganism host, but also promote the
XX folding of the fused insulin precursor, facilitate the solubility of the
XX fusion protein and decrease the intermolecular interactions among the

CC fusion proteins, thus allowing folding of the fused insulin precursor at
 CC commercially useful high concentrations. The procedural steps of cyanogen
 CC bromide cleavage, oxidative sulphitolytic and related purification steps
 CC can thus be eliminated, along with the use of high concentrations of
 CC mercapran or the use of hydrophobic absorbent resins

XX Sequence 86 AA:

Query Match 100.0%; Score 463; DB 2; Length 86;

Best Local Similarity 100.0%; Pred. No. 1.4e-43;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLYEALVYVCGERGFYTPKTRREAEADLVQGVLEGGPGAGSLQPLALEG 60
 DB 1 FVNQHLGSHLYEALVYVCGERGFYTPKTRREAEADLVQGVLEGGPGAGSLQPLALEG 60

QY 61 SLQKRGIVQCCTSTICSLYOLENYCN 86
 DB 61 SLQKRGIVQCCTSTICSLYOLENYCN 86

RESULT 4

AAB12770

XX AAB12770: standard; protein; 86 AA.

XX 22-NOV-2000 (first entry)

XX Human proinsulin protein sequence SEQ ID NO:2.

XX Human; insulin-like growth factor 1; IGF-1; proinsulin; insulin; mutant;
 XX variant; insulin-like growth factor binding protein; IGFBP-1; IGFBP-3;
 XX antidiabetic; neuroprotective; anorectic; tranquilliser; vulnerary;
 XX anorectic; cardiac; nephrotropic; dermatological; antihiv; antiviral;
 XX hyperglycaemia; obesity; lung disease; glomerulonephritis;
 XX interstitial nephritis; Turner's syndrome; Laron's syndrome;
 XX short stature; increased fat mass-to-lean ratio; immunological disorder;
 XX peripheral neuropathy; multiple sclerosis; muscular dystrophy;
 XX metabolic state; trauma; wounding; infection; HIV; skin disorder;
 XX human immunodeficiency virus; diabetes; heart dysfunction;
 XX kidney disorder; whole body growth disorder.

XX Homo sapiens.

XX MO200040612-A1.

XX 13-JUL-2000.

XX 05-JAN-2000; 2000MO-US000151.

XX 06-JAN-1999; 99US-0115010P.

XX (GETH) GENENTECH INC.

XX Dubaque Y, Lowman H;

XX WPI; 2000-465955/40.

XX Novel insulin-like growth factor (IGF) 1 mutants that selectively bind to
 XX IGF binding protein (IGFBP)-1 or IGFBP-3, used to improve the half-lives
 XX of IGF-1 and insulin.

XX Disclosure; Page 44; 48pp; English.

XX The present invention describes an insulin-like growth factor (IGF)-1
 XX variant (I), where an amino acid at position 3, 4, 5, 7, 10, 14, 17, 23,
 XX 24, 25, 43, 49 or 63, optionally in combination with an amino acid at
 XX position 12 and/or 16 of the native human IGF-1 sequence, is replaced
 XX with an alanine, glycine, or a serine residue. The residue at position 7
 XX may be replaced by any amino acid. (I) can have antidiabetic, cardiac,
 XX neuroprotective, anorectic, tranquiliser, vulnerary, anorectic,
 XX nephrotropic, dermatological, antihiv and antiviral activities. The IGF-1

CC mutants are used in any methods where IGFs or insulin are used, e.g. in
 CC treating hyperglycaemia, obesity-related, neurological, cardiac, renal,
 CC immunological, and anabolic disorders. These disorders include lung
 CC diseases, glomerulonephritis, interstitial nephritis, Turner's syndrome,
 CC Laron's syndrome, short stature, increased fat mass-to-lean ratios,
 CC immunological disorders, peripheral neuropathy, multiple sclerosis,
 CC muscular dystrophy, catabolic states, trauma, wounding, infection, human
 CC immunodeficiency virus (HIV), wounds, skin disorders, diabetes, heart
 CC dysfunction, kidney disorders, and whole body growth disorders. They can
 CC also be used for increasing serum and tissue levels of biologically active
 CC IGF or insulin in a mammal. The IGF-1 mutants improve the half-lives of IGF-
 CC 1 and insulin. The present sequence represents the native human
 CC proinsulin protein sequence, which is given in the exemplification of the
 CC present invention

XX Sequence 86 AA:

Query Match 100.0%; Score 463; DB 3; Length 86;

Best Local Similarity 100.0%; Pred. No. 1.4e-43;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLYEALVYVCGERGFYTPKTRREAEADLVQGVLEGGPGAGSLQPLALEG 60
 DB 1 FVNQHLGSHLYEALVYVCGERGFYTPKTRREAEADLVQGVLEGGPGAGSLQPLALEG 60

QY 61 SLQKRGIVQCCTSTICSLYOLENYCN 86
 DB 61 SLQKRGIVQCCTSTICSLYOLENYCN 86

RESULT 5

AAM48218

XX AAM48218: standard; protein; 86 AA.

XX 18-MAR-2002 (first entry)

XX Human proinsulin.

XX Antirheumatic; antiarthritic; osteopathic; cartilage disorder;
 XX insulin-like growth factor; IGF; binding protein; IGFBP;
 XX rheumatoid arthritis; osteoarthritis; proinsulin; human.

XX Homo sapiens.

XX MO200187323-A2.

XX 22-NOV-2001.

XX 16-MAY-2001; 2001MO-US015904.

XX 16-MAY-2000; 2000US-0204490P.

XX 15-NOV-2000; 2000US-0248985P.

XX (GETH) GENENTECH INC.

XX Dubaque Y, Filvaroff EH, Lowman HB;

XX WPI; 2002-082942/11.

XX Treating cartilage disorders including cartilage damage by injury or
 XX degenerative cartilaginous disorders, by contacting cartilage with
 XX insulin-like growth factor analog with altered affinity for IGF-binding
 XX proteins.

XX Disclosure; Fig 16; 136pp; English.

XX The present invention relates to a method for treating cartilage
 XX disorders. The method comprises contacting cartilage with an active agent
 XX such as insulin-like growth factor (IGF-1) analog with a binding affinity
 XX preference for IGF binding protein-3 (IGFBP-3) over IGFBP-1, an IGF-1
 XX analog with a binding affinity preference for IGFBP-1 over IGFBP-3, or a

CC IGFBP displacer peptide that prevents the interaction of IGF with an
CC IGFBP and does not bind to human IGF receptor. The method is useful for
CC treating cartilage disorders (CD), including degenerative CD, articular
CC CD such as rheumatoid arthritis and osteoarthritis. The present sequence
CC is human proinsulin, which was used to illustrate the invention
XX
SQ Sequence 86 AA:

Query Match 100.0%; Score 463; DB 5; Length 86;
Best Local Similarity 100.0%; Pred. No. 1,4e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHGCGSHLYEALYLVGSGRGGFYPTKTRREADLVGVGVELGGSPGAGSIQPLALEG 60
DB 2 FVNHGCGSHLYEALYLVGSGRGGFYPTKTRREADLVGVGVELGGSPGAGSIQPLALEG 60
QY 61 SLOKRGIVGOCCTSTICSLYQLENYCN 86
DB 61 SLOKRGIVGOCCTSTICSLYQLENYCN 86

RESULT 6

AD64463
ID AD64463 standard; protein; 86 AA.

AC AD64463;

DT 18-DEC-2003 (first entry)

DE Amino acid sequence for human proinsulin.

XX Immunassay; human C-peptide; HCP; immune complex; human; proinsulin.

XX Homo sapiens.

PN US2002160435-A1.

PD 31-OCT-2002.

PF 12-JUN-2001; 2001US-00878380.

PR 12-JUN-2000; 2000JP-00174691.

PA (KITA/) KITAJIMA S.

PA (KURA/) KURANO Y.

PA (NAKA/) NAKATSUBO K.

PA (NISH/) NISHIZONO I.

PI Kitajima S, Kurano Y, Nakatsubo K, Nishizono I;

DR WPI; 2003-765139/72.

PT Measuring human C-peptide, by reacting sample C-peptide with two

PT different human C-peptide antibodies that recognize different epitopes on

PT peptide, to form immune complex, separating and quantifying immune

PT complex.

PS Disclosure; SEQ ID NO 1; 20pp; English.

CC The present invention relates to an immunoassay for measuring human C-
CC peptide (HCP). The method comprises reacting HCP in a sample with a first
CC anti-HCP antibody and a second anti-HCP antibody which is immobilised on
CC a support, to form an immune complex, and separating and quantifying the
CC immune complex, where the first and second antibody recognises the
CC epitope existing in the region from 1-110 and 1-16 amino acid residues,
CC respectively, from the N-terminal end of HCP. Also disclosed is a kit for
CC measuring human C-peptide. The method is useful for measuring human C-
CC peptide. The method provides high reproducibility, high detection
CC sensitivity, and low cross-reactivity to proinsulin. The present sequence
CC represents the amino acid sequence for human proinsulin.

XX Sequence 86 AA;

Query Match 100.0%; Score 463; DB 7; Length 86;
Best Local Similarity 100.0%; Pred. No. 1,4e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHGCGSHLYEALYLVGSGRGGFYPTKTRREADLVGVGVELGGSPGAGSIQPLALEG 60
DB 1 FVNHGCGSHLYEALYLVGSGRGGFYPTKTRREADLVGVGVELGGSPGAGSIQPLALEG 60
QY 61 SLOKRGIVGOCCTSTICSLYQLENYCN 86
DB 61 SLOKRGIVGOCCTSTICSLYQLENYCN 86

RESULT 7

ADP16632
ID ADP16632 standard; protein; 86 AA.

AC ADP16632;

DT 12-FEB-2004 (first entry)

DE Human albumin fusion protein-related protein segid1734.

XX albumin fusion protein; albumin activity; human serum albumin;

XX serum osmotic pressure; shelf-life; stability; antidiabetic;

XX gene therapy; diabetes mellitus; human; gene; ds.

XX Homo sapiens.

PN WO2003060071-A2.

PD 24-JUL-2003.

PF 23-DEC-2002; 2002WO-US040891.

PR 21-DEC-2001; 2001US-0341811P.

PR 24-JAN-2002; 2002US-0350358P.

PR 28-JAN-2002; 2002US-0351360P.

PR 26-FEB-2002; 2002US-0359370P.

PR 28-FEB-2002; 2002US-0360000P.

PR 27-MAR-2002; 2002US-0367500P.

PR 08-APR-2002; 2002US-0370227P.

PR 10-MAY-2002; 2002US-0378950P.

PR 24-MAY-2002; 2002US-0382617P.

PR 28-MAY-2002; 2002US-0383123P.

PR 05-JUN-2002; 2002US-0385709P.

PR 10-JUL-2002; 2002US-0394625P.

PR 24-JUL-2002; 2002US-0398008P.

PR 09-AUG-2002; 2002US-0402131P.

PR 13-AUG-2002; 2002US-0402708P.

PR 18-SEP-2002; 2002US-0411355P.

PR 02-OCT-2002; 2002US-0414984P.

PR 11-OCT-2002; 2002US-0417611P.

PR 23-OCT-2002; 2002US-0420246P.

PR 05-NOV-2002; 2002US-0423623P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (DELZ) DELTA BIOTECHNOLOGY LTD.

PA (PRIN-) PRINCIPRIA PHARM CORP.

PI Ballance DJ, Turner AJ, Rosen CA, Haseltine WA,

DR WPI; 2003-598517/56.

DR N-PSDB; ADP16306.

PT New albumin fusion protein, useful for preparing a composition for

PT treating diabetes mellitus.

XX Example 4; SEQ ID NO 1734; 24pp; English.

XX This invention relates to a novel albumin fusion protein having albumin

CC or biological activity. Human serum albumin is responsible for a

CC significant proportion of the osmotic pressure of serum and also
CC functions as a carrier of endogenous and exogenous ligands. The fusion of
CC albumin to a therapeutic protein may increase shelf-life and stability of
CC the therapeutic protein. The albumin fusion protein of the invention may
CC allow production of compositions with antidiabetic activity whilst the
CC nucleotide sequence which encodes it may be useful for gene therapy. The
CC albumin fusion protein is useful for preparing a composition for treating
CC diabetes mellitus. The present sequence is that of a therapeutic protein
CC which was fused with human albumin to create a novel albumin fusion
CC protein of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at [ftp.wipo.int/pub/publishedpot_sequences](http://wipo.int/pub/publishedpot_sequences)
CC
XX
SQ Sequence 86 AA;

Query Match 100.0%; Score 463; DB 7; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.4e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLYEALTYVCGERGFYPTKTRREAEDELQGVGVLEGGPGAGSLQPLALEG 60
Db 1 FVNQHLGSHLYEALTYVCGERGFYPTKTRREAEDELQGVGVLEGGPGAGSLQPLALEG 60
QY 61 SLOKRGIVBOCCISICSLYOLENYCN 86
Db 61 SLOKRGIVBOCCISICSLYOLENYCN 86

RESULT 8
ADH21860
ID ADH21860 standard; protein; 86 AA.
XX
AC ADH21860;
XX
DT 11-MAR-2004 (first entry)
XX
DE Human long-acting insulin peptide, SEQ ID NO:657.
XX
XX Fusion protein; human serum albumin; HSA; therapeutic protein;
XX shelf-life; in vitro biological activity; in vivo biological activity;
XX metabolic disorder; endocrine disorder; diabetes; type 1; type 2;
XX diabetes-related condition; hyperglycaemia; neural disorder; neuropathy;
XX retinopathy; cardiovascular disorder; heart disease; renal disorder;
XX obesity; glucose level maintenance; weight loss; antidiabetic; cardiant;
XX anorectic; ophthalmologically; gene therapy.
XX
OS Homo sapiens.
XX
PN WO2003059934-A2.
XX
PD 24-JUL-2003.
XX
PF 23-DEC-2002; 2002WO-US040892.
XX
XX 21-DEC-2001; 2001US-0341811P.
XX 24-JAN-2002; 2002US-0350358P.
XX 26-FEB-2002; 2002US-0359370P.
XX 28-FEB-2002; 2002US-0360000P.
XX 27-MAR-2002; 2002US-0367500P.
XX 08-APR-2002; 2002US-0370227P.
XX 10-MAY-2002; 2002US-0378950P.
XX 24-JUL-2002; 2002US-0398008P.
XX 03-AUG-2002; 2002US-0402131P.
XX 13-AUG-2002; 2002US-0402708P.
XX 18-SEP-2002; 2002US-0411355P.
XX 02-OCT-2002; 2002US-0414984P.
XX 11-OCT-2002; 2002US-0417611P.
XX 23-OCT-2002; 2002US-0420246P.
XX 05-NOV-2002; 2002US-0423623P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX PA Rosen CA, Haseltine WA;
XX PI

XX WPI; 2003-596501/56.
XX DR N-PSDB; ADH21706.
XX
XX New albumin fusion protein, useful for preparing a composition for
XX treating diabetes mellitus.
XX
PS Disclosure; SEQ ID NO 657; 1086bp; English.
XX
XX The invention relates to fusion proteins comprising human serum albumin
XX (ADH2130) and a therapeutic polypeptide such as a therapeutic protein,
XX antibody or peptide or their variants or fragments. The therapeutic
XX protein may be fused to the N-terminus, the C-terminus or both termini of
XX albumin via a linker. The albumin component of the fusion proteins
XX protects the shelf-life and the in vitro and vivo biological activity of
XX the proteins compared with those of the corresponding therapeutic
XX proteins on their own. The invention also relates to nucleic acids
XX encoding albumin fusion proteins, vectors and host cells comprising an
XX albumin fusion protein nucleic acid, compositions and kits comprising an
XX albumin fusion protein. The method of extending the shelf-life of a
XX therapeutic protein by fusion with albumin, and the treatment of disease
XX using an albumin fusion protein. The albumin fusion proteins may be used
XX in the treatment of metabolic/endocrine disorders, diabetes and diabetes-
XX related conditions. Specifically the albumin fusion proteins may be used
XX to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
XX (especially neuropathy), retinopathy, cardiovascular disorders
XX (especially heart disease, renal disorders and obesity. The proteins may
XX also be used in a method of maintaining a basal glucose level in a
XX patient and in a method for losing weight. The present sequence is
XX related to the invention.
XX
SQ Sequence 86 AA;

Query Match 100.0%; Score 463; DB 7; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.4e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLYEALTYVCGERGFYPTKTRREAEDELQGVGVLEGGPGAGSLQPLALEG 60
Db 1 FVNQHLGSHLYEALTYVCGERGFYPTKTRREAEDELQGVGVLEGGPGAGSLQPLALEG 60
QY 61 SLOKRGIVBOCCISICSLYOLENYCN 86
Db 61 SLOKRGIVBOCCISICSLYOLENYCN 86

RESULT 9
AAP20036
ID AAP20036 standard; protein; 87 AA.
XX
AC AAP20036;
XX
DT 25-MAR-2003 (revised)
XX
DT 22-JUL-1992 (first entry)
XX
DE Human proinsulin.
XX
XX Proinsulin.
XX
OS Homo sapiens.
XX
PN EP55942-A.
XX
XX 14-JUL-1982.
XX
PD 31-DEC-1981; 81EP-00306190.
XX
PF 02-JAN-1981; 81US-00222010.
XX 23-JUL-1981; 81US-00286070.
XX 02-JAN-1982; 82US-00222010.
XX 03-MAR-1982; 82US-00354287.
XX
XX (UNVNY-) STATE UNIV NEW YORK.
XX PA
XX PI

```

XX
PI Inouye M, Nakamura K;
XX
DR WPI, 1982-59775E/29.
XX
DR N-PSDB; AAN20041.
XX
PT Plasmid cloning vehicles - useful for transforming bacterial hosts to
PT produce eukaryotic polypeptide(s).
XX
PS Disclosure; Fig 27; 11app; English.
XX
CC The sequence comprises human proinsulin. (Updated on 25-MAR-2003 to
CC correct PR field.)
XX
SQ Sequence 87 AA;
XX
Query Match 100.0%; Score 463; DB 1; Length 87;
Best Local Similarity 100.0%; Pred. No. 1,4e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 FVNQHLCGSHLVEAALYCGERGFYPTKTRREAEDELQGVQLGGGPGAGSIQPLALEG 60
Db 2 FVNQHLCGSHLVEAALYCGERGFYPTKTRREAEDELQGVQLGGGPGAGSIQPLALEG 61
QY 61 SLQKRGIVQCCCTSCISLYQLENYCN 86
Db 62 SLQKRGIVQCCCTSCISLYQLENYCN 87

RESULT 10
AAP40217
ID AAP40217 standard; protein; 87 AA.
XX
AC AAP40217;
XX
DT 25-MAR-2003 (revised)
DT 12-FEB-1992 (first entry)
XX
DE Sequence of the 32 N-terminal AAs of proinsulin.
XX
KM Hormone; cloning vector; phage resistant.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 2..31
FT /label= B-chain
FT Region 32..66
FT /label= C-chain
FT Region 67..87
FT /label= A-chain
XX
PN GB2126237-A.
XX
PD 21-MAR-1984.
XX
PF 01-SEP-1983; 83GB-00023468.
XX
PR 03-SEP-1982; 82US-00414280.
PR 05-SEP-1984; 84US-00647338.
XX
PA (ELIL ) LILLY & CO ELL.
XX
PI Hersherberg, CL; Rosteck PR;
XX
DR WPI; 1984-070793/12.
DR N-PSDB; AAN40179.
XX
PT Protecting bacteria from phage infection - by transformation with cloning
PT vector conng. segment with restriction and modification activity.
XX
PS Example; Fig 10; 28pp; English.
XX

```

```

CC Plasmid pTH alpha 1 was constructed by inserting a synthesised gene for
CC thymosin alpha 1 (AAN40178) into plasmid pBR322. It is used for the
CC construction of pTP24. The inventors claim a method for protecting
CC bacteria from phage infection - by transformation with cloning vector
CC conng. segment with restriction and modification activity. Prodn. of
CC plasmid pPR 26 or pPR27 which uses pTP24; and prodn. of plasmid pPR29
CC which uses a synthetic gene coding for the 32 N-terminal AAs of
CC proinsulin (see AAN40179). (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 87 AA;
XX
Query Match 100.0%; Score 463; DB 1; Length 87;
Best Local Similarity 100.0%; Pred. No. 1,4e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 FVNQHLCGSHLVEAALYCGERGFYPTKTRREAEDELQGVQLGGGPGAGSIQPLALEG 60
Db 2 FVNQHLCGSHLVEAALYCGERGFYPTKTRREAEDELQGVQLGGGPGAGSIQPLALEG 61
QY 61 SLQKRGIVQCCCTSCISLYQLENYCN 86
Db 62 SLQKRGIVQCCCTSCISLYQLENYCN 87

RESULT 11
AAP50127
ID AAP50127 standard; protein; 87 AA.
XX
AC AAP50127;
XX
DT 25-MAR-2003 (revised)
DT 16-AUG-2002 (revised)
DT 30-SEP-1991 (first entry)
XX
DE Sequence of the 32 N-terminal AAs of proinsulin.
XX
KM Selectable vector; autonomously replicating vector; expression vector.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 2..31
FT /label= A chain
FT Region 32..66
FT /label= B chain
FT Region 67..87
FT /label= A chain
XX
PN EP154539-A.
XX
PD 11-SEP-1985.
XX
PF 04-MAR-1985; 85EP-00301469.
XX
PR 06-MAR-1984; 84US-00586592.
XX
PA (ELIL ) LILLY & CO ELL.
XX
PI Schoner R, Schoner B;
XX
DR WPI; 1985-224921/37.
DR N-PSDB; AAN50152.
XX
PT New recombinant DNA expression vector - with autonomous replication and
PT on transcription generating polycistronic mna.
XX
PS Example; Fig 14; 118pp; English.
XX
CC The inventors claim a process for preparing selectable and autonomously
CC replicating recombinant DNA expression vectors which comprise 1) a
CC transcriptional and translational activating sequence which is in the
CC reading frame of a nucleotide sequence which codes for a peptide or

```

CC polypeptide; 2) a translational stop signal; 3) a translational start
 CC signal which is in the reading frame of a nucleotide sequence that codes
 CC for a functional polypeptide; and 4) an additional translational stop
 CC signal. The peptide or polypeptide coding sequence codes for 2-20 AAs,
 CC esp. AAs5012-550125. The functional polypeptide is esp. growth hormone,
 CC human insulin, interferon and human tissue plasminogen activator.
 CC (Updated on 16-AUG-2002 to add missing OS field.) (Updated on 25-MAR-2003
 CC to correct PA field.)
 CC
 XX
 SQ Sequence 87 AA;

Query Match 100.0%; Score 463; DB 1; Length 87;
 Best Local Similarity 100.0%; Pred. No. 1.4e-43;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHGCGSHLVEALYVCGERGFYTPKTRRAEDLVQGVLELGGPGAGSLQPLALEG 60
 DB 2 FVNHGCGSHLVEALYVCGERGFYTPKTRRAEDLVQGVLELGGPGAGSLQPLALEG 61
 QY 61 SLQKRGIVEOCCTSGICSLYLENYCN 86
 DB 62 SLQKRGIVEOCCTSGICSLYLENYCN 87

RESULT 12

AAP50060
 ID AAP50060 standard; protein; 87 AA.

AC AAP50060;
 DT 25-MAR-2003 (revised)
 DT 16-AUG-2002 (revised)
 DT 11-NOV-1991 (first entry)

XX
 DE Synthetic proinsulin.

XX
 KW Proinsulin; vector; proteinaceous granule.

XX
 OS Homo sapiens.

XX
 FH Key Location/Qualifiers
 FT Region 1..30
 FT Region /label= B chain.
 FT Region 31..65
 FT Region /label= C chain.
 FT Region 66..86
 FT Region /label= A chain.

XX
 PN EPI59123-A.
 XX
 PD 23-OCT-1985.
 XX
 PF 04-MAR-1985; 85EP-00301468.
 XX
 PR 06-MAR-1984; 84US-00586582.
 PR 26-JUL-1984; 84US-00634920.
 PR 31-JAN-1985; 85US-00697090.

XX
 PA (BLIL) LILLY & CO ELI.
 XX
 PI Hsiung HM, Schoner RG, Schoner BE;
 XX
 DR WPI; 1985-265090/43.
 DR N-PSDB; AAN50082.

XX
 PT New selectable and autonomously replicating DNA expression vector -
 PT useful in producing proteinaceous granules in cell transformants, esp.
 PT for prodn. of bovine growth hormone derivs.
 XX
 PX Disclosure; Fig 14, 115pp; English.

XX
 CC The synthetic proinsulin gene is expressed in a new selectable and
 CC autonomously replicating recombinant DNA expression vector comprising a

CC runaway replicon and a transcriptional and translational activating
 CC sequence in the reading frame of the proinsulin coding sequence; the
 CC sequence contg. a translational stop signal. Host cells contg. the
 CC vector, which is esp. plasmid pc2103, are cultured, and proinsulin is
 CC produced as a highly homogeneous species of proteinaceous granule. The
 CC granule can be readily isolated from cell lysates and is stable on
 CC washing with urea or detergent solns. at low concns. The granule contains
 CC at least 50% of proinsulin and all isolation operations are simplified.
 CC (Updated on 16-AUG-2002 to add missing OS field.) (Updated on 25-MAR-2003
 CC to correct PA field.)
 CC
 XX
 SQ Sequence 87 AA;

Query Match 100.0%; Score 463; DB 1; Length 87;
 Best Local Similarity 100.0%; Pred. No. 1.4e-43;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHGCGSHLVEALYVCGERGFYTPKTRRAEDLVQGVLELGGPGAGSLQPLALEG 60
 DB 2 FVNHGCGSHLVEALYVCGERGFYTPKTRRAEDLVQGVLELGGPGAGSLQPLALEG 61
 QY 61 SLQKRGIVEOCCTSGICSLYLENYCN 86
 DB 62 SLQKRGIVEOCCTSGICSLYLENYCN 87

RESULT 13

AAP61090
 ID AAP61090 standard; protein; 87 AA.

AC AAP61090;
 DT 28-FEB-1992 (first entry)

XX
 DE Sequence encoded by the structural gene for human proinsulin.

XX
 KW Recombinant plasmid; E.coli expression vector; secretion vector.

XX
 OS Homo sapiens.

XX
 PN US4624926-A.

XX
 PD 25-NOV-1986.

XX
 PF 03-MAR-1982; 82US-00354287.

XX
 PR 02-JAN-1981; 81US-00222010.
 PR 23-JUL-1981; 81US-00286070.

XX
 PA (UNIV-) UNIV OF NEW YORK.

XX
 PI Inouye M, Nakamura K;

XX
 DR WPI; 1986-331802/50.
 DR N-PSDB; AAN60872.

XX
 PT New recombinant plasmid(s) - contg. DNA sequences encoding exogenous
 PT polypeptide and outer membrane protein of E coli.
 XX
 PX Example; Fig 27; 44pp; English.

XX
 CC The inventors claim new recombinant plasmids contg. a DNA sequence
 CC encoding a polypeptide, which is foreign to E.coli, in reading phase with
 CC a DNA SQ, coding for at least one functional fragment derived from an
 CC outer membrane lipoprotein gene of E.coli. The foreign gene may be for
 CC human insulin. The lipoprotein gene functional fragment may be the
 CC promoter, the 5'-UTR, the 3'-UTR or the transcription termination signal
 CC provided that it includes at least the promoter
 XX
 SQ Sequence 87 AA;

XX
 CC Query Match 100.0%; Score 463; DB 1; Length 87;
 CC Best Local Similarity 100.0%; Pred. No. 1.4e-43;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNOHLCGSHLVEALYIVCGERGFYPTKTRRAEDLVQVGVLGSGPGAGSLQPLALEG 60
DB 2 FVNOHLCGSHLVEALYIVCGERGFYPTKTRRAEDLVQVGVLGSGPGAGSLQPLALEG 61
QY 61 SLOKRGIVGCCCTSIICSLYOLENYCN 86
DB 62 SLOKRGIVGCCCTSIICSLYOLENYCN 87

RESULT 14

AAK32367
ID AAK32367 standard; protein; 87 AA.

AC AAK32367;
DT 25-MAR-2003 (revised)
DT 18-JUN-1993 (first entry)
DE Proinsulin protein sequence.

KM Human; proinsulin; vector; pUC19; pINS; CAT; pUC-CAT-proinsulin;
KW Insulin analogue; type I; type II; diabetes.

OS Synthetic.

PN WO9303174-A1.

PD 18-FEB-1993.

PF 31-JUL-1992; 92MO-US006451.

PR 08-AUG-1991; 91US-00741938.

PR 30-JUL-1992; 92US-00918953.

PA (SCIO-) SCIOS INC.
(Pfizer) PFIZER INC.

PI Andy RJ, Larson ER;

DR WPI; 1993-076530/09.

DR N-PSDB; AAQ37003.

PT New hepato selective and peripheral selective human insulin analogues -
PT and their corresp. DNA, for treatment of type I and type II diabetes.

PS Disclosure; Fig 2b; 58pp; English.

CC This sequence represents human proinsulin and was decoded from the
CC sequences given in AAQ36996-7001. The cDNA fragment coding for proinsulin
CC was inserted into plasmid vector pUC19 and digested with KpnI and
CC HindIII. This resulted in the formation of the vector pINS. A fragment
CC encoding amino acids 1-73 of CAT (see AAQ37002) was inserted into pINS
CC to give a plasmid which contained DNA sequences which coded for amino
CC acids 1-73 of CAT, an 8 amino acid linker sequence and human proinsulin.
CC This plasmid, pUC-CAT-proinsulin, could be used in the formation of
CC insulin analogues which may be used in the treatment of types I and II
CC diabetes. (Updated on 25-MAR-2003 to correct PN field.)

QY Sequence 87 AA;

Query Match 100.0%; Score 463; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 1.4e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNOHLCGSHLVEALYIVCGERGFYPTKTRRAEDLVQVGVLGSGPGAGSLQPLALEG 60
DB 2 FVNOHLCGSHLVEALYIVCGERGFYPTKTRRAEDLVQVGVLGSGPGAGSLQPLALEG 61
QY 61 SLOKRGIVGCCCTSIICSLYOLENYCN 86
DB 62 SLOKRGIVGCCCTSIICSLYOLENYCN 87

RESULT 15

AAK07682
ID AAK07682 standard; protein; 88 AA.

AC AAK07682;
DT 25-MAR-2003 (revised)
DT 09-JAN-2003 (revised)
DT 13-FEB-1991 (first entry)
DE Modified human insulin precursor.

KM Human insulin precursor; cathepsin C.

OS Homo sapiens.

FN Key Location/Qualifiers

FT Peptide 1..2 /label= N-terminal initiating dipeptide

FT Peptide 3..32 /label= native human insulin B-chain

FT Peptide 33..67 /label= natural connecting peptide of human proinsulin

FT Peptide 68..88 /label= native human insulin A-chain

PN EP397420-A.

PD 14-NOV-1990.

PF 04-MAY-1990; 90EP-00304890.

PR 09-MAY-1989; 89US-00349472.

PA (BLIL) LILLY & CO ELLI.

PI Becker GW, Furman TC, Mackellar WC, McDonough JP;

DR WPI; 1990-343372/46.

PT Human insulin precursor - contg. Met-Tyr or Met-Arg initiating dipeptide
PT for controlled removal by cathepsin C.

PS Disclosure; Page 3; 8pp; English.

CC This modified human insulin precursor comprises an N-terminal initiating
CC dipeptide, chosen from Met-Tyr or Met-Arg, which does not define a
CC cathepsin C dipeptide removal stop point. This dipeptide is linked to the
CC natural human insulin B-chain, natural human proinsulin connecting
CC peptide and natural human insulin A-chain. Dipeptide removal is
CC carefully controlled to obtain the desired prod. without further
CC degradation occurring, irrespective of whether the next dipeptide in the
CC sequence defines a cathepsin C stop point. (Updated on 09-JAN-2003 to add
CC missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)

QY Sequence 88 AA;

Query Match 100.0%; Score 463; DB 2; Length 88;
Best Local Similarity 100.0%; Pred. No. 1.5e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNOHLCGSHLVEALYIVCGERGFYPTKTRRAEDLVQVGVLGSGPGAGSLQPLALEG 60
DB 3 FVNOHLCGSHLVEALYIVCGERGFYPTKTRRAEDLVQVGVLGSGPGAGSLQPLALEG 62
QY 61 SLOKRGIVGCCCTSIICSLYOLENYCN 86
DB 63 SLOKRGIVGCCCTSIICSLYOLENYCN 88

Search completed: November 2, 2004, 20:11:46

Sat Nov 6 18:59:24 2004

Job time : 86.9852 secs

us-10-054-873-4.rag

Page 9

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 2, 2004, 20:02:41 [Search time 20.9446 Seconds
(without alignments)
272.306 Million cell updates/sec]

Title: US-10-054-873-4

Sequence: 1 FVNHLCGSHLVEALYLVCGERGFFYPTKTRDAEDLVGVDELGGGPGAGSLPFLAEFG 86

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: Issued Patents, AA:*

1: /cgn2_6/ptodata/1/aa/5A.COMB.pep:*

2: /cgn2_6/ptodata/1/aa/5B.COMB.pep:*

3: /cgn2_6/ptodata/1/aa/6A.COMB.pep:*

4: /cgn2_6/ptodata/1/aa/6B.COMB.pep:*

5: /cgn2_6/ptodata/1/aa/PCPTS.COMB.pep:*

6: /cgn2_6/ptodata/1/aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	463	100.0	86	4	US-09-477-924-2
2	463	100.0	86	4	US-09-723-981-2
3	463	100.0	86	4	US-09-723-896-2
4	463	100.0	86	4	US-09-878-380-1
5	463	100.0	96	3	US-09-134-836-4
6	463	100.0	96	3	US-09-386-303A-4
7	463	100.0	96	4	US-09-947-553-4
8	463	100.0	97	1	US-08-160-376A-4
9	463	100.0	110	3	US-08-950-720A-11
10	463	100.0	110	3	US-08-589-028-2
11	463	100.0	110	3	US-08-784-582-2
12	463	100.0	110	3	US-08-785-271-2
13	463	100.0	110	4	US-08-472-701-2
14	463	100.0	110	4	US-09-185-852-2
15	463	100.0	110	4	US-09-815-229-3
16	463	100.0	110	4	US-09-617-389B-20
17	463	100.0	110	4	US-09-323-738-2
18	463	100.0	110	4	US-09-015-359-7
19	463	100.0	117	5	PCR-US95-08596-2
20	463	100.0	117	4	US-09-280-030-63
21	463	100.0	130	4	US-09-280-030-62
22	463	100.0	151	2	US-08-508-664-15
23	463	100.0	167	2	US-08-508-664-16
24	463	100.0	167	1	US-07-918-993-8
25	463	100.0	167	1	US-08-081-661-8
26	457	98.7	96	2	US-09-134-836-8
27	457	98.7	96	3	US-09-386-303A-5

28	457	98.7	96	4	US-09-947-553-5	Sequence 5, Appli
29	457	98.7	97	1	US-08-389-487-7	Sequence 7, Appli
30	456	98.5	90	1	US-08-030-731A-43	Sequence 43, Appli
31	456	98.5	98	4	US-09-701-968-7	Sequence 7, Appli
32	456	98.5	99	4	US-09-701-968-8	Sequence 8, Appli
33	456	98.5	100	4	US-09-701-968-9	Sequence 9, Appli
34	449	97.0	110	4	US-09-574-443-1	Sequence 1, Appli
35	446	96.3	97	3	US-09-099-307-6	Sequence 6, Appli
36	444	95.7	97	3	US-09-099-307-8	Sequence 8, Appli
37	443	95.7	110	3	US-08-589-028-4	Sequence 4, Appli
38	443	95.7	110	3	US-08-784-582-4	Sequence 4, Appli
39	443	95.7	110	3	US-08-785-271-4	Sequence 7, Appli
40	440	95.0	97	3	US-09-099-307-7	Sequence 7, Appli
41	435	94.0	97	3	US-09-099-307-11	Sequence 11, Appli
42	398	86.0	91	4	US-09-676-787-7	Sequence 7, Appli
43	300	64.8	56	4	US-09-815-229-10	Sequence 10, Appli
44	297.5	63.2	67	3	US-08-981-988A-1	Sequence 1, Appli
45	290.5	62.7	83	3	US-08-981-988A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-477-924-2
; Sequence 2, Application US/09477924
; Patent No. 6403764
; GENERAL INFORMATION:
; APPLICANT: Dubaque, Yves
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: P1712R1-1
; CURRENT APPLICATION NUMBER: US/09/477,924
; CURRENT FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 2
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-477-924-2

Query Match 100.0%; Score 463; DB 4; Length 86;
Best Local Similarity 100.0%; Pred. No. 2.2e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	FVNHLCGSHLVEALYLVCGERGFFYPTKTRDAEDLVGVDELGGGPGAGSLPFLAEFG	60
DB	1	FVNHLCGSHLVEALYLVCGERGFFYPTKTRDAEDLVGVDELGGGPGAGSLPFLAEFG	60
QY	61	SLQKRGIVEOCCTCSICSLYOLENYCN	86
DB	61	SLQKRGIVEOCCTCSICSLYOLENYCN	86

RESULT 2
US-09-723-981-2
; Sequence 2, Application US/09723981
; Patent No. 6506874
; GENERAL INFORMATION:
; APPLICANT: Dubaque, Yves
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: P1712R1
; CURRENT APPLICATION NUMBER: US/09/723,981
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/477,923
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 2
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-723-981-2

Query Match 100.0%; Score 463; DB 4; Length 86;

Best Local Similarity 100.0%; Pred. No. 2.2e-47; Indels 0; Gaps 0;

Matches 86; Conservative 0; Mismatches 0;

QY 1 FVNOHLCGSHLYEALYVCGERGFYPTKTRREABDLQVGVLEGGPGAGSLQPLALEG 60

DB 1 FVNOHLCGSHLYEALYVCGERGFYPTKTRREABDLQVGVLEGGPGAGSLQPLALEG 60

QY 61 SLOKRGIVEOCCTSIQSLYQLENYCN 86

DB 61 SLOKRGIVEOCCTSIQSLYQLENYCN 86

RESULT 3

US-09-723-896-2

Sequence 2, Application US/09723896

Patent No. 6509443

GENERAL INFORMATION:

APPLICANT: Dubaque, Yves

TITLE OF INVENTION: PROTEIN VARIANTS

FILE REFERENCE: P1712R1

CURRENT APPLICATION NUMBER: US/09/723,896

CURRENT FILING DATE: 2000-11-28

PRIOR APPLICATION NUMBER: US/09/477,923

NUMBER OF SEQ ID NOS: 6

SEQ ID NO 2

LENGTH: 86

TYPE: PRT

ORGANISM: Homo sapiens

US-09-723-896-2

Query Match

Best Local Similarity 100.0%; Score 463; DB 4; Length 86;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNOHLCGSHLYEALYVCGERGFYPTKTRREABDLQVGVLEGGPGAGSLQPLALEG 60

DB 1 FVNOHLCGSHLYEALYVCGERGFYPTKTRREABDLQVGVLEGGPGAGSLQPLALEG 60

QY 61 SLOKRGIVEOCCTSIQSLYQLENYCN 86

DB 61 SLOKRGIVEOCCTSIQSLYQLENYCN 86

RESULT 4

US-09-878-380-1

Sequence 1, Application US/09878380

Patent No. 6534281

GENERAL INFORMATION:

APPLICANT: Fujirebio Inc.

APPLICANT: KITAJIMA, Sachiko

APPLICANT: KURANO, Yoshihiro

APPLICANT: NAKATSUBO, Kaoru

APPLICANT: NISHIZONO, Isao

TITLE OF INVENTION: Immunoassay for Measuring Human C-Peptide and Kit therefor

FILE REFERENCE: 0760-0291P

CURRENT APPLICATION NUMBER: US/09/878,380

CURRENT FILING DATE: 2001-06-12

PRIOR APPLICATION NUMBER: JP 2000-174691

PRIOR FILING DATE: 2000-06-12

NUMBER OF SEQ ID NOS: 2

SOFTWARE: Patent version 3.1

SEQ ID NO 1

LENGTH: 86

TYPE: PRT

ORGANISM: Homo sapiens

US-09-878-380-1

Query Match

Best Local Similarity 100.0%; Score 463; DB 4; Length 86;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNOHLCGSHLYEALYVCGERGFYPTKTRREABDLQVGVLEGGPGAGSLQPLALEG 60

DB 1 FVNOHLCGSHLYEALYVCGERGFYPTKTRREABDLQVGVLEGGPGAGSLQPLALEG 60

QY 61 SLOKRGIVEOCCTSIQSLYQLENYCN 86

DB 61 SLOKRGIVEOCCTSIQSLYQLENYCN 86

RESULT 5

US-09-134-836-4

Sequence 4, Application US/09134836

Patent No. 5986048

GENERAL INFORMATION:

APPLICANT: Rubroder, Franz-Josef

APPLICANT: Keller, Reinhold

TITLE OF INVENTION: Improved process for obtaining

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &

STREET: 1300 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/134,836

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Leslie McDaniel

REGISTRATION NUMBER: 34,872

REFERENCE/DOCKET NUMBER: 02481.1600-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 408-4000

TELEFAX: (202) 408-4400

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 96 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: Escherichia coli

FEATURE:

NAME/KEY: Protein

LOCATION: 1..96

US-09-134-836-4

Query Match

Best Local Similarity 100.0%; Score 463; DB 2; Length 96;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNOHLCGSHLYEALYVCGERGFYPTKTRREABDLQVGVLEGGPGAGSLQPLALEG 60

DB 1 FVNOHLCGSHLYEALYVCGERGFYPTKTRREABDLQVGVLEGGPGAGSLQPLALEG 60

QY 61 SLOKRGIVEOCCTSIQSLYQLENYCN 86

DB 61 SLOKRGIVEOCCTSIQSLYQLENYCN 86

RESULT 6

US-09-134-836-4

Sequence 4, Application US/09134836

Patent No. 5986048

GENERAL INFORMATION:

APPLICANT: Rubroder, Franz-Josef

APPLICANT: Keller, Reinhold

TITLE OF INVENTION: Improved process for obtaining

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &

STREET: 1300 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/134,836

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Leslie McDaniel

REGISTRATION NUMBER: 34,872

REFERENCE/DOCKET NUMBER: 02481.1600-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 408-4000

TELEFAX: (202) 408-4400

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 96 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: Escherichia coli

FEATURE:

NAME/KEY: Protein

LOCATION: 1..96

US-09-134-836-4

Query Match

Best Local Similarity 100.0%; Score 463; DB 2; Length 96;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNOHLCGSHLYEALYVCGERGFYPTKTRREABDLQVGVLEGGPGAGSLQPLALEG 60

DB 1 FVNOHLCGSHLYEALYVCGERGFYPTKTRREABDLQVGVLEGGPGAGSLQPLALEG 60

QY 61 SLOKRGIVEOCCTSIQSLYQLENYCN 86

DB 61 SLOKRGIVEOCCTSIQSLYQLENYCN 86

US-09-386-303A-4
Sequence 4, Application US/09386303A
Patent No. 6380355
GENERAL INFORMATION:
APPLICANT: Rubroder, Franz-Josef
Keller, Reinhold
TITLE OF INVENTION: Improved process for obtaining
insulin precursors having correctly bonded cystine bridges
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farrabow, Garrett &
Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/386,303A
FILING DATE: 31-Aug-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/134,836
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Leslie McDona11
REGISTRATION NUMBER: 34,872
REFERENCE/DOCKET NUMBER: 02481.1600-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 96 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: Protein
LOCATION: 1..96
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-386-303A-4

Query Match 100.0%; Score 463; DB 3; Length 96;
Best Local Similarity 100.0%; Pred. No. 2.6e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNOHLCGSHLVEALYLVCGERGFFYTPKTRREADLDVGVGVGGSPGAGSLQPLALEG 60
DB 11 FVNOHLCGSHLVEALYLVCGERGFFYTPKTRREADLDVGVGVGGSPGAGSLQPLALEG 70
QY 61 SLQKRGIVGECCTSTICSLVQLENYCN 86
DB 71 SLQKRGIVGECCTSTICSLVQLENYCN 96

RESULT 7
US-09-947-563-4
Sequence 4, Application US/09947563
Patent No. 6727346
GENERAL INFORMATION:
APPLICANT: Rubroder, Franz-Josef
Keller, Reinhold
TITLE OF INVENTION: Improved process for obtaining
insulin precursors having correctly bonded cystine bridges

NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farrabow, Garrett &
Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/947,563
FILING DATE: 07-Sep-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/134,836
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Leslie McDona11
REGISTRATION NUMBER: 34,872
REFERENCE/DOCKET NUMBER: 02481.1600-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 96 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: Protein
LOCATION: 1..96
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-947-563-4

Query Match 100.0%; Score 463; DB 4; Length 96;
Best Local Similarity 100.0%; Pred. No. 2.6e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNOHLCGSHLVEALYLVCGERGFFYTPKTRREADLDVGVGVGGSPGAGSLQPLALEG 60
DB 11 FVNOHLCGSHLVEALYLVCGERGFFYTPKTRREADLDVGVGVGGSPGAGSLQPLALEG 70
QY 61 SLQKRGIVGECCTSTICSLVQLENYCN 86
DB 71 SLQKRGIVGECCTSTICSLVQLENYCN 96

RESULT 8
US-08-160-376A-4
Sequence 4, Application US/08160376A
Patent No. 5473049
GENERAL INFORMATION:
APPLICANT: Obermeier, Ranier
APPLICANT: Geil, Martin
APPLICANT: Ludwig, Jurgen
APPLICANT: Sabel, Walter
TITLE OF INVENTION: Process For Obtaining Proinsulin
Possessing Correctly Linked
Cysteine Bridges
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenneth A. Genoni, Esq.
STREET: Rt. 202-206 No. 5473049th/P.O. Box 2500
CITY: Somerville

STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 08876-1258
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
COMPUTER: IBM 386
OPERATING SYSTEM: WINDOWS 3.1
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/160,376A
FILING DATE: December 1, 1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB P 4240420.7
FILING DATE: December 2, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Barbara V. Maurer, Esq.
REGISTRATION NUMBER: 31,287
REFERENCE/DOCKET NUMBER: HOB 92/F 384
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 231-4079
TELEFAX: (908) 231-2255
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 97 Amino Acids
TYPE: Amino Acid (AA)
TOPOLOGY: not relevant
US-08-160-376A-4

Query Match 100.0%; Score 463; DB 1; Length 97;
Best Local Similarity 100.0%; Pred. No. 2.6e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHGCGSHLYEALYVCGERGFFYPKTRREADLDQVGVLEGGPGAGSLQPLALEG 60
DB 12 FVNHGCGSHLYEALYVCGERGFFYPKTRREADLDQVGVLEGGPGAGSLQPLALEG 71
QY 61 SLQKRGIVECCCTSIQSLYLENYCN 86
DB 72 SLQKRGIVECCCTSIQSLYLENYCN 97

RESULT 9
US-08-950-720A-11
Sequence 11, Application US/08950720A
Patent No. 6046028
GENERAL INFORMATION:
APPLICANT: Conklin, Darrell C.
APPLICANT: Lofton-Day, Catherine E.
APPLICANT: Lok, Si
APPLICANT: Jaspers, Stephen R.
TITLE OF INVENTION: INSULIN HOMOLOG
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Zymogenetics, Inc.
STREET: 1201 Eastlake Avenue East
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98102
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/950,720A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: Sawislak, Deborah A
REGISTRATION NUMBER: 37,438
REFERENCE/DOCKET NUMBER: 96-09
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-442-6672
TELEFAX: 206-442-6678
TELEX:
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 110 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6046028e
US-08-950-720A-11

Query Match 100.0%; Score 463; DB 3; Length 110;
Best Local Similarity 100.0%; Pred. No. 3e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHGCGSHLYEALYVCGERGFFYPKTRREADLDQVGVLEGGPGAGSLQPLALEG 60
DB 25 FVNHGCGSHLYEALYVCGERGFFYPKTRREADLDQVGVLEGGPGAGSLQPLALEG 84
QY 61 SLQKRGIVECCCTSIQSLYLENYCN 86
DB 85 SLQKRGIVECCCTSIQSLYLENYCN 110

RESULT 10
US-08-589-028-2
Sequence 2, Application US/08589028
Patent No. 6087129
GENERAL INFORMATION:
APPLICANT: Newgard, Christopher B.
APPLICANT: Halban, Philippe
APPLICANT: No. 6087129minston, Karl D.
APPLICANT: Clark, Samuel A.
APPLICANT: Thigpen, Anice E.
APPLICANT: Quade, Christian
APPLICANT: Kruse, Fred
TITLE OF INVENTION: Recombinant Expression of Proteins From
NUMBER OF SEQUENCES: 50
TITLE OF INVENTION: Secretory Cell Lines
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P. O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/589,028
FILING DATE: Concurrently Herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 47,642
REFERENCE/DOCKET NUMBER: UTSD:426\HYL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 474-7577
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 110 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear

US-08-589-028-2

Query Match 100.0%; Score 463; DB 3; Length 110;
Best Local Similarity 100.0%; Pred. No. 3e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
DB 25 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84
QY 61 SLQKRGIVECCCTSCISLYOLENYCN 86
DB 85 SLQKRGIVECCCTSCISLYOLENYCN 110

RESULT 11

US-08-784-582-2

; Sequence 2, Application US/08784582

; Patent No. 610707

; GENERAL INFORMATION:

; APPLICANT: Newgard, Christopher B.

; APPLICANT: Halban, Philippe A.

; APPLICANT: No. 610707mington, Karl D.

; APPLICANT: Clark, Samuel A.

; APPLICANT: Thigpen, Anice E.

; APPLICANT: Quade, Christian

; APPLICANT: Kruse, Fred

; TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM

; NUMBER OF INVENTION: SECRETORY CELL LINES

; NUMBER OF SEQUENCES: 79

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: Arnold, White & Durkee

; STREET: P. O. Box 4433

; CITY: Houston

; STATE: Texas

; COUNTRY: USA

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/784,582

; FILING DATE: Concurrently Herewith

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/028,427

; FILING DATE: 15-OCT-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/589,028

; FILING DATE: 19-JAN-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Highlander, Steven L.

; REGISTRATION NUMBER: 37,642

; REFERENCE/DOCKET NUMBER: UTSD:514

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 512/474-7577

; TELEFAX: 512/474-7577

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 110 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

US-08-784-582-2

DB 25 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84
QY 61 SLQKRGIVECCCTSCISLYOLENYCN 86
DB 85 SLQKRGIVECCCTSCISLYOLENYCN 110

RESULT 12

US-08-785-271-2

; Sequence 2, Application US/08785271

; Patent No. 6194176

; GENERAL INFORMATION:

; APPLICANT: Newgard, Christopher B.

; APPLICANT: Halban, Philippe A.

; APPLICANT: No. 6194176mington, Karl D.

; APPLICANT: Clark, Samuel A.

; APPLICANT: Thigpen, Anice E.

; APPLICANT: Quade, Christian

; APPLICANT: Kruse, Fred

; TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM

; NUMBER OF INVENTION: SECRETORY CELL LINES

; NUMBER OF SEQUENCES: 56

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: Arnold, White & Durkee

; STREET: P. O. Box 4433

; CITY: Houston

; STATE: Texas

; COUNTRY: USA

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/785,271

; FILING DATE: Concurrently Herewith

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/589,028

; FILING DATE: 19-JAN-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Highlander, Steven L.

; REGISTRATION NUMBER: 37,642

; REFERENCE/DOCKET NUMBER: UTSD:513

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 512/474-7577

; TELEFAX: 512/474-7577

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 110 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

US-08-785-271-2

Query Match 100.0%; Score 463; DB 3; Length 110;
Best Local Similarity 100.0%; Pred. No. 3e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
DB 25 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84
QY 61 SLQKRGIVECCCTSCISLYOLENYCN 86
DB 85 SLQKRGIVECCCTSCISLYOLENYCN 110

RESULT 13

US-08-472-701-2

; Sequence 2, Application US/08472701

Patent No. 6509165
GENERAL INFORMATION:
APPLICANT: Griffin, Ann C.
APPLICANT: Hickey, William F.
TITLE OF INVENTION: Detection and Treatment Methods for
TITLE OF INVENTION: Type 1 Diabetes
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,701
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/272,220
FILING DATE: 08-JULY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A., Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: DCI-092DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 110 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-701-2

Query Match 100.0%; Score 463; DB 4; Length 110;
Best Local Similarity 100.0%; Pred. No. 3e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLYEALYVCGERGFYTPKTRRAEDLVQVGLGGPGAGSLQPLALEG 60
DB 25 FVNQHLGSHLYEALYVCGERGFYTPKTRRAEDLVQVGLGGPGAGSLQPLALEG 84

QY 61 SLQKRGIVEQCCTSGISLYOLENYCN 86
DB 85 SLQKRGIVEQCCTSGISLYOLENYCN 110

RESULT 14
US-09-185-852-2
Sequence 2, Application US/09185852
Patent No. 6537806
GENERAL INFORMATION:
APPLICANT: Osborne, William R.A.
APPLICANT: Ramesh, Nagaraajan
TITLE OF INVENTION: Compositions and Methods for Treating Diabetes
FILE REFERENCE: P-UM 3264
CURRENT APPLICATION NUMBER: US/09/185,852
CURRENT FILING DATE: 1998-11-04
EARLIER APPLICATION NUMBER: 60/087,660
EARLIER FILING DATE: 1998-06-02
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 110
TYPE: PRT

ORGANISM: Homo sapiens
US-09-185-852-2

Query Match 100.0%; Score 463; DB 4; Length 110;
Best Local Similarity 100.0%; Pred. No. 3e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLYEALYVCGERGFYTPKTRRAEDLVQVGLGGPGAGSLQPLALEG 60
DB 25 FVNQHLGSHLYEALYVCGERGFYTPKTRRAEDLVQVGLGGPGAGSLQPLALEG 84

QY 61 SLQKRGIVEQCCTSGISLYOLENYCN 86
DB 85 SLQKRGIVEQCCTSGISLYOLENYCN 110

RESULT 15
US-09-815-229-3
Sequence 3, Application US/09815229
Patent No. 6689747
GENERAL INFORMATION:
APPLICANT: Filvaroff, Ellen H.
APPLICANT: Okumu, Franklin W.
TITLE OF INVENTION: USE OF INSULIN FOR THE TREATMENT OF CARTILLAGENOUS DISORDERS
FILE REFERENCE: P178681US
CURRENT APPLICATION NUMBER: US/09/815,229
CURRENT FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: US 60/192,103
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 17
SEQ ID NO 3
LENGTH: 110
TYPE: PRT
ORGANISM: Homo sapiens
US-09-815-229-3

Query Match 100.0%; Score 463; DB 4; Length 110;
Best Local Similarity 100.0%; Pred. No. 3e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLYEALYVCGERGFYTPKTRRAEDLVQVGLGGPGAGSLQPLALEG 60
DB 25 FVNQHLGSHLYEALYVCGERGFYTPKTRRAEDLVQVGLGGPGAGSLQPLALEG 84

QY 61 SLQKRGIVEQCCTSGISLYOLENYCN 86
DB 85 SLQKRGIVEQCCTSGISLYOLENYCN 110

Search completed: November 2, 2004, 20:24:34
Job time: 21.9446 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 2, 2004, 19:59:41 : Search time 15.8672 Seconds
(without alignments)
521.495 Million cell updates/sec

Title: US-10-054-873-4

Perfect score: 463

Sequence: 1 FVNOHLCGSHVLEALVWCG.....IVEOCCTCSICSLYQENYCN 86

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Lasting first 45 summaries

Database:

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	463	100.0	110	1	IPHU
2	463	100.0	110	2	A42179
3	456	98.5	110	2	B42179
4	456	98.5	110	2	J00178
5	424	91.6	110	1	INRB
6	417	90.1	110	1	IPDG
7	394	85.1	110	1	IPHO
8	394	85.1	110	1	INMS2
9	394	85.1	110	1	IPRT2
10	392	84.7	110	2	A39863
11	392	84.7	110	2	A39863
12	385	83.2	110	1	IPRT1
13	383	82.7	110	1	IPPG
14	366.5	79.2	105	1	IPBO
15	366	79.0	108	1	INMS1
16	334.5	72.2	108	2	S09278
17	320.5	69.2	77	1	INSH
18	314	67.8	110	1	IPGR
19	277.5	59.9	109	1	IPRTDU
20	276.5	59.7	103	2	A51221
21	265.5	57.3	106	1	IPX12
22	265.5	57.3	107	1	IPCH
23	262.5	56.7	106	1	IPX11
24	256.5	55.4	51	1	INEL
25	256.5	55.4	51	1	INMHF
26	256.5	55.4	51	1	INMHP
27	256.5	55.4	51	1	IPDK
28	256	55.3	96	2	PC7082
29	254.5	55.0	51	1	INHY

30	251.5	54.3	51	1	INMSP
31	250.5	54.1	51	2	A5151
32	246.5	53.2	51	1	INMVA
33	246.5	53.2	51	1	INCT
34	246.5	53.2	51	1	INMHL
35	245.5	53.0	51	1	INCT
36	244.5	52.8	51	1	INMSO
37	239.5	51.7	51	2	J00362
38	234.5	50.6	51	1	INCB
39	231.5	50.0	51	1	INCS
40	227.5	49.1	51	1	INOS
41	227.5	49.1	51	1	INTR
42	227.5	49.1	51	1	A61129
43	227.5	49.1	51	1	INPO
44	227.5	49.1	51	2	A60414
45	225	48.6	52	2	S44470

ALIGNMENTS

RESULT 1

IPHU

insulin precursor [validated] - human

N/Alternate names: preproinsulin

C/Species: Homo sapiens (man)

C/Date: 23-Oct-1981 #sequence revision 23-Oct-1981 #text change 09-Jul-2004

C/Accession: A93222; A94253; A93216; A94251; A93144; A92075; A91186; A91114; A01579; SE

R/Bell, G.I.; Pictet, R.U.; Rutter, W.J.; Cordell, B.; Tischer, E.; Goodman, H.M.

Nature 284, 26-32, 1980

A/Title: Sequence of the human insulin gene.

A/Reference number: A93222; PMID:80120725; PMID:6243748

A/Accession: A93222

A/Molecule type: DNA

A/Residues: 1-110 <BEU>

A/Cross-references: UNIPROT:P01308; GB:J00265; NID:g186429; PIDN:AAA59172.1; PID:g38682

R/Bell, G.I.; Swain, W.F.; Pictet, R.; Cordell, B.; Goodman, H.M.; Rutter, W.J.

Nature 282, 525-527, 1979

A/Title: Nucleotide sequence of a cDNA clone encoding human preproinsulin.

A/Reference number: A93216; PMID:80054779; PMID:503234

A/Accession: A93216

A/Molecule type: mRNA

A/Residues: 1-110 <BEU>

A/Cross-references: GB:J00265; NID:g186429; PIDN:AAA59172.1; PID:g386828

R/Sures, I.; Goeddel, D.V.; Gray, A.; Ullrich, A.

Science 208, 57-59, 1980

A/Title: Nucleotide sequence of human preproinsulin complementary DNA.

A/Reference number: A94251; PMID:80147417; PMID:6927840

A/Accession: A94251

A/Molecule type: mRNA

A/Residues: 1-110 <SUR>

A/Cross-references: GB:J00265; NID:g186429; PIDN:AAA59172.1; PID:g386828

R/Nicol, D.S.H.W.; Smith, L.F.

Nature 187, 483-485, 1960

A/Title: Amino acid sequence of human insulin.

A/Reference number: A93144

A/Accession: A93144

A/Molecule type: protein

A/Residues: 25-54/90-110 <NIC>

R/Oyer, P.E.; Cho, S.; Peterson, J.D.; Steiner, D.F.

J. Biol. Chem. 246, 1375-1386, 1971

A/Title: Studies on human proinsulin. Isolation and amino acid sequence of the human part

A/Reference number: A92075; PMID:71116410; PMID:5101771

A/Accession: A92075

A/Molecule type: protein

A:Residues: 57-87 <ORF>
R:Ko, A.; Smyth, D.G.; Markussen, J.; Sundby, F.
Eur. J. Biochem. 20, 190-199, 1971
A:Title: Amino acid sequence of the C-peptide of human proinsulin.
A:Reference number: A91186; MUID:71257722; PMID:5560404
A:Accession: A91186
A:Molecule type: protein
A:Residues: 57-87 <KOA>
R:Knudsen, A.M.; Julier, C.; Beressi, J.P.; Bolteard, C.; Froquel, P.; Lathrop, M.; Bell
Nucleic Acids Res. 4, 105-110, 1993
A:Title: Susceptibility to insulin dependent diabetes mellitus maps to a 4.1 kb segment
A:Reference number: 158114; MUID:93364428; PMID:8358440
A:Accession: 158114
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-59,63-110 <RES>
R:Sieber, P.; Kamber, B.; Hartmann, A.; Joehl, A.; Rinkler, B.; Rittel, W.
Helv. Chim. Acta 57, 2617-2621, 1974
A:Title: Totalsynthese von Humaninsulin unter gezielter Bildung der Disulfidbindungen.
A:Reference number: A91636; MUID:75077277; PMID:4443293
A:Contents: annotation; synthesis
A>Note: disulfide-bonded human insulin was synthesized; the synthetic hormone was identical
A>Note: article in German with English abstract
R:Nathani, V.K.
Hoppe-Seyler's Z. Physiol. Chem. 354, 659-672, 1973
A:Title: The synthesis of C-peptide of human proinsulin.
A:Reference number: A91658; MUID:75040007; PMID:4803504
A:Contents: annotation; synthesis of residues 57-87
R:Geiger, R.; Jaeger, G.; Koenig, W.
Chem. Ber. 106, 2347-2352, 1973
A:Title: Synthesis of the complete sequence of human proinsulin C-peptide and its [Glu-9
A:Reference number: A90914
A:Contents: annotation; synthesis of residues 57-87
R:Kaufmann, J.E.; Imringer, J.C.; Halban, P.A.
Biochem. J. 310, 869-874, 1995
A:Title: Sequence requirements for proinsulin processing at the B-chain/C-peptide junction
A:Reference number: S58661; MUID:96033185; PMID:7575420
A:Contents: annotation; site-directed mutagenesis study of proteolytic processing
A:Genetics:
A:Gene: GDB:INS
A:Cross-references: GDB:119349; OMIM:176730
A:Map position: 11p15.5-11p15.5
A:Introns: 63/1
A:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-54/Domain: insulin chain B #status experimental <BCH>
F:125-87/Domain: insulin chain A #status experimental <MAT>
F:157-87/Domain: connecting C peptide #status experimental <CH>
F:90-110/Domain: insulin chain A #status experimental <ACH>
F:31-96,43-109,95-100/Disulfide bonds: #status experimental
Query Match 100.0%; Score 463; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 6,8e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FVNQHLGSHLVKLVYVCGERGFYTPKTRREADLDVGVGLGGGAGSLQPLALEG 60
DB 25 FVNQHLGSHLVKLVYVCGERGFYTPKTRREADLDVGVGLGGGAGSLQPLALEG 84
QY 61 SLQKRGIVEOCTCSISLYOLENYCN 86
DB 85 SLQKRGIVEOCTCSISLYOLENYCN 110
RESULT 2
Insulin precursor - chimpanzee
A:Species: Pan troglodytes (chimpanzee)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A42179; S22058
R:Reino, S.; Bell, G.I.; Li, W.H.

Mol. Biol. Evol. 9, 193-203, 1992
A:Title: Sequences of primate insulin genes support the hypothesis of a slower rate of
A:Reference number: A42179; MUID:92219953; PMID:1560757
A:Accession: A42179
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-110 <SET>
A:Cross-references: UNIPROT:P30410; EMBL:X61089; NID:938251; PIDN:CAA43403.1; PID:938252
A>Note: sequence extracted from NCBI backbone (NCBI:P:95067)
A:Genetics:
A:Introns: 63/1
A:Superfamily: insulin
Query Match 100.0%; Score 463; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 6,8e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FVNQHLGSHLVKLVYVCGERGFYTPKTRREADLDVGVGLGGGAGSLQPLALEG 60
DB 25 FVNQHLGSHLVKLVYVCGERGFYTPKTRREADLDVGVGLGGGAGSLQPLALEG 84
QY 61 SLQKRGIVEOCTCSISLYOLENYCN 86
DB 85 SLQKRGIVEOCTCSISLYOLENYCN 110
RESULT 3
Insulin precursor - green monkey
A:Species: Cercopithecus aethiops (green monkey, grivet)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A42179; A052332; S16494; S22058
R:Reino, S.; Bell, G.I.; Li, W.H.
Mol. Biol. Evol. 9, 193-203, 1992
A:Title: Sequences of primate insulin genes support the hypothesis of a slower rate of
A:Reference number: A42179; MUID:92219953; PMID:1560757
A:Accession: A42179
A:Molecule type: DNA
A:Residues: 1-110 <SET>
A:Cross-references: UNIPROT:P30407; EMBL:X61092; NID:922808; PIDN:CAA43405.1; PID:92280
A>Note: sequence extracted from NCBI backbone (NCBI:95185; NCBI:P:95194)
R:Reino, S.; Bell, G.I.; Li, W.H.
J. Biol. Chem. 247, 4866-4871, 1972
A:Title: Determination of the amino acid sequence of the monkey, sheep, and dog proinsu
A:Reference number: A92111; MUID:72258016; PMID:4626369
A:Accession: A05232
A:Molecule type: protein
A:Residues: 57-87 <PER>
A:Genetics:
A:Gene: GDB:INS
A:Cross-references: GDB:119349; OMIM:176730
A:Map position: 11p15.5-11p15.5
A:Introns: 63/1
A:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-54/Domain: insulin chain B #status experimental <BCH>
F:125-87/Domain: insulin chain A #status experimental <MAT>
F:157-87/Domain: connecting peptide #status experimental <CH>
F:90-110/Domain: insulin chain A #status predicted <ACH>
F:31-96,43-109,95-100/Disulfide bonds: #status predicted
Query Match 98.8%; Score 456; DB 2; Length 110;
Best Local Similarity 98.8%; Pred. No. 3,9e-42;
Matches 85; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 FVNQHLGSHLVKLVYVCGERGFYTPKTRREADLDVGVGLGGGAGSLQPLALEG 60
DB 25 FVNQHLGSHLVKLVYVCGERGFYTPKTRREADLDVGVGLGGGAGSLQPLALEG 84
QY 61 SLQKRGIVEOCTCSISLYOLENYCN 86
DB 85 SLQKRGIVEOCTCSISLYOLENYCN 110
RESULT 4

J00178
insulin precursor - crab-eating macaque
C:Species: Macaca fascicularis (crab-eating macaque)
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C:Accession: J00178
R:Wetzel, W.; Gronenberg, J.; Leineweber, M.; Wengenmayer, F.; Winnacker, E.L.
Gene 19, 179-183, 1992
A:Title: The nucleotide sequence of cDNA coding for preproinsulin from the primate Macaca fascicularis
A:Reference number: J00178; MUID:83080474; PMID:6184262
A:Accession: J00178
A:Molecule type: mRNA
A:Residues: 1-110 <MUT>
A:Cross-references: UNIPROT:P30406; GB:J00336; NID:g342121; PIDN:AAA6849.1; PID:g342122
C:Superfamily: insulin
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-54/Domain: insulin chain B #status predicted <BCH>
F:55-89/Domain: insulin connecting C peptide #status predicted <CPT>
F:90-110/Domain: insulin chain A #status predicted <ACH>
F:31-96,43-109,95-100/Disulfide bonds: #status predicted

Query Match 98.5% Score 456; DB 2; Length 110;
Best Local Similarity 98.8% Pred. No. 3,9e-42;
Matches 85; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYPTKTRREAEDLQVQVELGGGPGASGLQPLALEG 60
DB 25 FVNQHLGSHLVEALYLVCGERGFFYPTKTRREAEDPQVQVELGGGPGASGLQPLALEG 84
QY 61 SLOKRGIVQCCTSTICSLYLENYCN 86
DB 85 SLOKRGIVQCCTSTICSLYLENYCN 110

RESULT 5

INR8
insulin precursor - rabbit
N:Alternate names: preproinsulin
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 12-Apr-1964 #sequence_revision 23-Aug-1997 #text_change 09-Jul-2004
C:Accession: A53438; A01581
R:Devaskar, S.V.; Giddings, S.V.; Rajakumar, P.A.; Carnaghi, L.R.; Menon, R.K.; Zahm, D.
J. Biol. Chem. 269, 8445-8454, 1994
A:Title: Insulin gene expression and insulin synthesis in mammalian neuronal cells.
A:Reference number: A53438; MUID:94179230; PMID:8132571
A:Accession: A53438
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-110 <MUT>
A:Cross-references: UNIPROT:P01311; GB:U03610; NID:g467970; PIDN:AAA19033.1; PID:g467971
R:Smith, L.F.
Am. J. Med. 40, 662-666, 1966
A:Title: Species variation in the amino acid sequence of insulin.
A:Reference number: A90029; MUID:66160119; PMID:5949593
A:Accession: A01581
A:Molecule type: protein
A:Residues: 25-54,90-110 <SMI>
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-54/Domain: insulin chain B #status experimental <BCH>
F:55-89/Domain: insulin connecting C peptide #status experimental <CPT>
F:90-110/Domain: insulin chain A #status experimental <ACH>
F:31-96,43-109,95-100/Disulfide bonds: #status predicted

Query Match 91.6% Score 424; DB 1; Length 110;
Best Local Similarity 90.7% Pred. No. 1.1e-35;
Matches 78; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYPTKTRREAEDLQVQVELGGGPGASGLQPLALEG 60
DB 25 FVNQHLGSHLVEALYLVCGERGFFYPTKTRREAEDLQVQVELGGGPGASGLQPLALEG 84

QY 61 SLOKRGIVQCCTSTICSLYLENYCN 86
DB 85 SLOKRGIVQCCTSTICSLYLENYCN 110

RESULT 6

IPDG
insulin precursor - dog
C:Species: Canis lupus familiaris (dog)
C>Date: 24-Apr-1984 #sequence_revision 15-Nov-1984 #text_change 09-Jul-2004
C:Accession: A92413; A01587; S16493
R:Kwok, S.C.M.; Chan, S.J.; Steiner, D.F.
J. Biol. Chem. 258, 2357-2363, 1983
A:Title: Cloning and nucleotide sequence analysis of the dog insulin gene. Coded amino acid sequence
A:Reference number: A92413; MUID:83109071; PMID:6286142
A:Accession: A92413
A:Molecule type: DNA
A:Residues: 1-110 <SMI>
A:Cross-references: UNIPROT:P01321; GB:V00179; GB:U00442; NID:g9994; PIDN:CAA23475.1; P
R:Smith, L.F.
Am. J. Med. 40, 662-666, 1966
A:Title: Species variation in the amino acid sequence of insulin.
A:Reference number: A90029; MUID:66160119; PMID:5949593
A:Accession: A01587
A:Molecule type: protein
A:Residues: 25-54,90-110 <SMI>
R:Peterson, J.D.; Nehrllich, S.; Oyer, P.E.; Steiner, D.F.
J. Biol. Chem. 247, 4866-4871, 1972
A:Title: Determination of the amino acid sequence of the monkey, sheep, and dog proinsulin
A:Reference number: A92111; MUID:72258016; PMID:4626369
A:Accession: S16493

Query Match 90.1% Score 417; DB 1; Length 110;
Best Local Similarity 89.5% Pred. No. 6.3e-38;
Matches 77; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYPTKTRREAEDLQVQVELGGGPGASGLQPLALEG 60
DB 25 FVNQHLGSHLVEALYLVCGERGFFYPTKTRREAEDLQVQVELGGGPGASGLQPLALEG 84
QY 61 SLOKRGIVQCCTSTICSLYLENYCN 86
DB 85 SLOKRGIVQCCTSTICSLYLENYCN 110

RESULT 7

IPHO
insulin precursor - horse
C:Species: Equus caballus (domestic horse)
C>Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 09-Jul-2004
C:Accession: A01580; A92120
R:Harris, J.I.; Sanger, F.; Naughton, M.A.
Arch. Biochem. Biophys. 65, 427-428, 1956
A:Title: Species differences in insulin.
A:Reference number: A90082
A:Accession: A01580
A:Molecule type: protein
A:Residues: 1-30,65-86 <HAR>
A:Cross-references: UNIPROT:P01310
R:Tagger, H.S.; Steiner, D.F.
J. Biol. Chem. 247, 7936-7940, 1972
A:Title: Primary structures of the proinsulin connecting peptides of the rat and horse

A:Reference number: A92120; MUID:73061498; PMID:4640931
 A:Accession: A92120
 A:Molecule type: protein
 A:Residues: 33-63 <TAG>
 C:Comment: X's at positions 31-32 and 64-65 represent paired basic residues assumed (by
 C:Superfamily: Insulin
 C:Keywords: hormone; pancreas
 F:1-30/Domain: Insulin chain B #status experimental <BCH>
 F:31-63/Domain: Insulin chain B #status experimental <MAT>
 F:33-63/Domain: connecting peptide #status experimental <CPEP>
 F:66-86/Domain: Insulin chain A #status experimental <ACH>
 F:72,73,74,75,76/Disulfide bonds: #status predicted

Query Match 85.1%; Score 394; DB 1; Length 86;
 Best Local Similarity 84.9%; Pred. No. 1.5e-35;
 Matches 73; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEALYVCGERGFYPTKTRPADLDQVQLVGGGPGAGSLDPLALEG 60
 DB 1 FVNHLCGSHLVEALYVCGERGFYPTKTRPADLDQVQLVGGGPGAGSLDPLALEG 60
 61 SLOKRGIVDCCSTICSLYLENYCN 86
 DB 61 PQXKGIVDCCSTICSLYLENYCN 86

RESULT 8
 INNS2
 Insulin 2 precursor - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 31-Mar-1992 #sequence revision 14-Jul-1994 #text_change 09-Jul-2004
 C:Accession: A26342; B48172; A61012; B01592
 R:Wentworth, B.M.; Schaefer, I.M.; Valla-Komaroff, L.; Chitgwin, J.M.
 J. Mol. Evol. 23, 305-312, 1986
 A:Title: Characterization of the two nonallelic genes encoding mouse preproinsulin.
 A:Reference number: A92965; MUID:87169768; PMID:3104603
 A:Accession: A26342
 A:Molecule type: DNA
 A:Residues: 1-110 <MEN>
 A:Cross-references: UNIPROT:P01326; GB:X04724; NID:952714; PIDN:CAA28433.1; PID:952715
 R:Sawa, T.; Ohgaki, S.; Morioka, H.; Yano, S.
 J. Mol. Endocrinol. 5, 61-67, 1990
 A:Title: Molecular cloning and DNA sequence analysis of preproinsulin genes in the NON
 A:Reference number: A48172; MUID:90372969; PMID:2397023
 A:Accession: B48172
 A:Molecule type: DNA
 A:Status: not compared with conceptual translation
 A:Residues: 1-110 <SAM>
 R:Linder, S.; Nielsen, J.H.; Hansen, B.; Wellinder, B.S.
 J. Chromatogr. 462, 243-254, 1989
 A:Title: Reversed-phase high-performance liquid chromatographic analyses of insulin bios
 A:Reference number: A61012; MUID:8922078; PMID:2651585
 A:Accession: A61012
 A:Molecule type: protein
 A:Residues: 57-87 <LIN>
 R:Buenzli, H.F.; Glattchar, B.; Kunz, P.; Muehlaupt, E.; Humbel, R.E.
 Hoppe-Seyler's Z. Physiol. Chem. 353, 451-458, 1972
 A:Title: Amino acid sequence of the two insulins from mouse (Mus musculus).
 A:Reference number: A01592; MUID:72189455; PMID:5063718
 A:Accession: B01592
 A:Molecule type: protein
 A:Residues: 25-54,90-110 <BUB>
 C:Genetics:
 A:Introns: 63/1
 C:Superfamily: Insulin
 C:Keywords: hormone; pancreas
 F:1-54/Domain: Insulin chain B #status predicted <SIG>
 F:25-54/Domain: Insulin chain B #status experimental <BCH>
 F:25-54,90-110/Domain: Insulin chain B #status experimental <MAT>
 F:57-87/Domain: connecting peptide #status experimental <CPEP>
 F:90-110/Domain: Insulin chain A #status experimental <ACH>
 F:31-96,43-109,95-100/Disulfide bonds: #status predicted

Query Match 85.1%; Score 394; DB 1; Length 110;
 Best Local Similarity 84.9%; Pred. No. 1.9e-35;
 Matches 73; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEALYVCGERGFYPTKTRPADLDQVQLVGGGPGAGSLDPLALEG 60
 DB 25 FVNHLCGSHLVEALYVCGERGFYPTKTRPADLDQVQLVGGGPGAGSLDPLALEG 84
 61 SLOKRGIVDCCSTICSLYLENYCN 86
 DB 85 AQKRGIVDCCSTICSLYLENYCN 110

RESULT 9
 IPR2
 Insulin 2 precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 23-Oct-1981 #sequence revision 23-Oct-1981 #text_change 09-Jul-2004
 C:Accession: B90789; B94231; C92120; I64880; A01590; B92120
 R:Umedico, P.; Rosenthal, N.; Estratadias, A.; Gilbert, W.; Kolodner, R.; Tizard, R.
 Cell 18, 545-558, 1979
 A:Title: The structure and evolution of the two nonallelic rat preproinsulin genes.
 A:Reference number: A90789; MUID:80045035; PMID:498284
 A:Accession: B90789
 A:Molecule type: DNA
 A:Residues: 1-110 <LOM>
 A:Cross-references: UNIPROT:P01323; GB:J00748; NID:9204958; PIDN:AAA41443.1; PID:9204958
 R:Steiner, D.F.; Clark, J.L.; Nolan, C.; Rubenstein, A.H.; Margoliah, E.; Aten, B.; Oy.
 Recent Prog. Horm. Res. 25, 207-282, 1969
 A:Title: Proinsulin and the biosynthesis of insulin.
 A:Reference number: A94231; MUID:70067613; PMID:4311938
 A:Accession: B94231
 A:Molecule type: protein
 A:Residues: 25-54,90-110 <STB>
 R:Tagger, H.S.; Steiner, D.F.
 J. Biol. Chem. 247, 7936-7940, 1972
 A:Title: Primary structures of the proinsulin connecting peptides of the rat and horse.
 A:Reference number: A92120; MUID:73061498; PMID:4640931
 A:Accession: C92120
 A:Molecule type: protein
 A:Residues: 57-87 <TAG>
 R:Umedico, P.T.; Rosenthal, N.; Kolodner, R.; Estratadias, A.; Gilbert, W.
 Ann. N. Y. Acad. Sci. 343, 425-432, 1980
 A:Title: The structure of rat preproinsulin genes.
 A:Reference number: I51945; MUID:80240379; PMID:6249167
 A:Accession: I64880
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-110 <RGS>
 A:Cross-references: GB:M25585; NID:9204950; PIDN:AAA41440.1; PID:9204952
 C:Genetics:
 A:Gene: INS2
 A:Introns: 63/1
 C:Superfamily: Insulin
 C:Keywords: hormone; pancreas
 F:1-24/Domain: Insulin chain B #status predicted <SIG>
 F:25-54/Domain: Insulin chain B #status experimental <BCH>
 F:25-54,90-110/Domain: Insulin chain B #status experimental <MAT>
 F:57-87/Domain: connecting peptide #status experimental <CPEP>
 F:90-110/Domain: Insulin chain A #status experimental <ACH>
 F:31-96,43-109,95-100/Disulfide bonds: #status experimental

Query Match 85.1%; Score 394; DB 1; Length 110;
 Best Local Similarity 84.9%; Pred. No. 1.9e-35;
 Matches 73; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEALYVCGERGFYPTKTRPADLDQVQLVGGGPGAGSLDPLALEG 60
 DB 25 FVNHLCGSHLVEALYVCGERGFYPTKTRPADLDQVQLVGGGPGAGSLDPLALEG 84
 61 SLOKRGIVDCCSTICSLYLENYCN 86
 DB 85 AQKRGIVDCCSTICSLYLENYCN 110

```
RESULT 10
A:Accession: A39883
A:Species: Aotus trivirgatus (douroucoulis, night monkey, owl monkey)
C:Date: 27-Nov-1991 #sequence_revision 27-Nov-1991 #text_change 09-Jul-2004
C:Accession: A39883
R:Seino, S.; Steiner, D.F.; Bell, G.I.
Proc. Natl. Acad. Sci. U.S.A. 84, 7423-7427, 1987
A:Title: Sequence of a New World primate insulin having low biological potency and immu-
A:Reference number: A39883; PMID:8604119; PMID:3118367
A:Accession: A39883
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-108 <SEI>
A:Cross-references: UNIPROT:P10604; GB:J02989; NID:g176555; PIDN:AAA35374.1; PID:g176556
C:Superfamily: insulin

Query Match
Best Local Similarity 84.7%; Score 392; DB 2; Length 108;
Matches 73; Conservative 4; Mismatches 7; Indels 2; Gaps 1;

QY 1 FVNQHLCSHLYEALVYVCGRGFFYPTKTRREAEDELQVGVGPGAGSLQPLALEG 60
DB 25 FVNQHLCSHLYEALVYVCGRGFFYPTKTRREAEDELQVGVGPGAGSLQPLALEG 82
QY 61 SLOKRGIVQCCCTSCISLYOLENYCN 86
DB 83 PMQKRGIVDQCCCTSCISLYOLENYCN 108

RESULT 11
A:Accession: A39883
A:Species: Aotus trivirgatus (douroucoulis, night monkey, owl monkey)
C:Date: 27-Nov-1991 #sequence_revision 27-Nov-1991 #text_change 09-Jul-2004
C:Accession: A39883
R:Seino, S.; Steiner, D.F.; Bell, G.I.
Proc. Natl. Acad. Sci. U.S.A. 84, 7423-7427, 1987
A:Title: Sequence of a New World primate insulin having low biological potency and immu-
A:Reference number: A39883; PMID:8604119; PMID:3118367
A:Accession: A39883
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-108 <SEI>
A:Cross-references: UNIPROT:P10604; GB:J02989; NID:g176555; PIDN:AAA35374.1; PID:g176556
C:Superfamily: insulin

Query Match
Best Local Similarity 84.7%; Score 392; DB 2; Length 110;
Matches 73; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 FVNQHLCSHLYEALVYVCGRGFFYPTKTRREAEDELQVGVGPGAGSLQPLALEG 60
DB 25 FVNQHLCSHLYEALVYVCGRGFFYPTKTRREAEDELQVGVGPGAGSLQPLALEG 84
QY 61 SLOKRGIVQCCCTSCISLYOLENYCN 86
DB 85 AQKRGIVDQCCCTSCISLYOLENYCN 110

RESULT 12
A:Accession: A39788
A:Species: Rattus norvegicus (Norway rat)
C:Date: 23-Oct-1981 #sequence_revision 23-Oct-1981 #text_change 09-Jul-2004
C:Accession: A39788; A90789; A94231; B92120; I51945; A01589
R:Corbelli, B.; Bell, G.; Tischer, E.; Deheto, F.M.; Ullrich, A.; Picot, R.; Rutter, W.J.
Cell 18, 533-543, 1979
A:Title: Isolation and characterization of a cloned rat insulin gene.
A:Reference number: A90788; PMID:80045034; PMID:498283
A:Accession: A90788

A:Molecule type: DNA
A:Residues: 1-110 <COR>
A:Cross-references: UNIPROT:P01322; GB:J00747; NID:g204956; PIDN:AAA41442.1; PID:g2045
R:Lomedico, P.; Rosenthal, N.; Estratidis, A.; Gilbert, W.; Kolodner, R.; Tizard, R.
Cell 18, 545-558, 1979
A:Title: The structure and evolution of the two nonallelic rat preproinsulin genes.
A:Reference number: A90789; PMID:80045035; PMID:498284
A:Accession: A90789
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-110 <DOM>
A:Cross-references: GB:J00747; NID:g204956; PIDN:AAA41442.1; PID:g204957
R:Steiner, D.F.; Clark, U.V.; Nollan, C.; Rubenstein, A.H.; Margolis, E.; Aten, B.; O
Recent Prog. Horm. Res. 25, 207-282, 1969
A:Title: Proinsulin and the biosynthesis of insulin.
A:Reference number: A94231; PMID:70067613; PMID:4311938
A:Accession: A94231
A:Molecule type: protein
A:Residues: 25-54-90-110 <STB>
R:Tagger, H.S.; Steiner, D.F.
U. Biol. Chem. 247, 7936-7940, 1972
A:Title: Primary structures of the proinsulin connecting peptides of the rat and horse
A:Reference number: A92120; PMID:73061498; PMID:4640931
A:Accession: B92120
A:Molecule type: protein
A:Residues: 57-87 <TAG>
R:Lomedico, P.T.; Rosenthal, N.; Kolodner, R.; Estratidis, A.; Gilbert, W.
Ann. N. Y. Acad. Sci. 343, 425-442, 1980
A:Title: The structure of rat preproinsulin genes.
A:Reference number: I51945; PMID:80240379; PMID:6249167
A:Accession: I51945
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-110 <RES>
A:Cross-references: GB:M25584; NID:g204947; PIDN:AAA41439.1; PID:g204948
C:Genetics:
A:Gene: INS1
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-54/Domain: insulin chain B #status experimental <SCH>
F:57-87/Domain: product: insulin #status experimental <MAT>
F:90-110/Domain: connecting peptide #status experimental <CEP>
F:90-110/Domain: insulin chain A #status experimental <ACH>
F:31-96/43-109,95-100/Disulfide bonds: #status experimental

Query Match
Best Local Similarity 83.2%; Score 385; DB 1; Length 110;
Matches 72; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 FVNQHLCSHLYEALVYVCGRGFFYPTKTRREAEDELQVGVGPGAGSLQPLALEG 60
DB 25 FVNQHLCSHLYEALVYVCGRGFFYPTKTRREAEDELQVGVGPGAGSLQPLALEG 84
QY 61 SLOKRGIVQCCCTSCISLYOLENYCN 86
DB 85 ARQKRGIVDQCCCTSCISLYOLENYCN 110

RESULT 13
A:Accession: A01583
A:Species: Sus scrofa domestica (domestic pig)
C:Date: 22-Jun-1981 #sequence_revision 22-Jun-1981 #text_change 16-Jul-1999
C:Accession: A01583; A94572; S16492; A60835; B60835
R:Chance, R.E.; Ellis, R.M.; Bromer, W.W.
Science 161, 165-167, 1968
A:Title: Porcine proinsulin: characterization and amino acid sequence.
A:Reference number: A94240; PMID:6828685; PMID:5657063
A:Accession: A01583
A:Molecule type: protein
A:Residues: 1-34; Q, 36-84 <CHA>
R:Chance, R.E.
submitted to the Atlas, July 1970
```

A:Reference number: A94572
A:Accession: A94572
A:Molecule type: protein
A:Residues: 1-84 <CH2>
R:Brown, H.; Sanger, F.; Kitai, R.
Biochem. J. 60, 556-565, 1955
A:Title: The structure of pig and sheep insulins.
A:Reference number: A90344
A:Accession: S16492
A:Molecule type: protein
A:Residues: 1-30;31-51 <BRO>
R:Strel, L.; Damgaard, U.
Horm. Metab. Res. 20, 476-480, 1988
A:Title: Proinsulin heterogeneity in pigs.
A:Reference number: A60835; PMID:89032178; PMID:3181865
A:Accession: A60835
A:Molecule type: protein
A:Residues: 33-38,40-62 <SNE>
A:Note: the authors report the characterization of a connecting peptide variant lacking
A:Accession: B60835
A:Molecule type: protein
A:Residues: 33-62 <SN2>
R:Blundell, T.; Dodson, G.; Hodgkin, D.; Mercola, D.
Adv. Protein Chem. 26, 279-402, 1972
A:Title: Insulin: the structure in the crystal and its reflection in chemistry and biol
A:Reference number: A90017
A:Contents: annotation; X-ray crystallography, 1.9 angstroms
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-30/Domain: insulin chain B #status experimental <BCH>
F:33-63/Domain: connecting peptide #status experimental <MAT>
F:64-84/Domain: insulin chain A #status experimental <ACH>
F:7-10,19-63,69-74/Disulfide bonds: #status experimental

Query Match 82.7%; Score 383; DB 1; Length 84;
Best Local Similarity 86.0%; Pred. No. 2,3e-34;
Matches 74; Conservative 1; Mismatches 9; Indels 2; Gaps 1;

QY 1 FVNQHLGSHLYALTYVCGERGFYTPKRRRAEDLYQVGLGGPGAGSLQPLALEG 60
DB 1 FVNQHLGSHLYALTYVCGERGFYTPKRRRAEDLYQVGLGGPGAGSLQPLALEG 58

QY 61 SLOKRGIVQCCISGLYOLENYCN 86
DB 59 PPOKRGIVQCCISGLYOLENYCN 84

RESULT 14
IPBO
insulin precursor - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 24-Apr-1984 #sequence revision 22-Apr-1995 #next change 09-Jul-2004
C:Accession: A40909; A92080; A92074; A91185; A90342; A90341; S48184; S48185; S46258; A01
R:ID Agostino, U.; Young, M.A.; White, J.W.; Besch, F.K.; Field, U.B.; Frazer, M.L.
Mol. Endocrinol. 1, 327-331, 1987
A:Title: Cloning and nucleotide sequence analysis of complementary deoxyribonucleic acid
A:Reference number: A40909; PMID:88288209; PMID:2456452
A:Accession: A40909
A:Molecule type: mRNA
A:Residues: 1-105 <DAA>
A:Cross-references: UNIPROT:P01317; GB:MS4579; NID:G163578; PTD:AAA30722.1; PID:G163579
A:Experimental source: fetal pancreas
R:Noian, C.; Margolis, E.; Peterson, J.D.; Steiner, D.F.
J. Biol. Chem. 246, 2780-2795, 1971
A:Title: The structure of bovine proinsulin.
A:Reference number: A92080; PMID:71166442; PMID:4928892
A:Accession: A92080
A:Molecule type: protein
A:Residues: 25-105 <NOL>
R:Steiner, D.F.; Cho, S.; Over, P.E.; Terris, S.; Peterson, J.D.; Rubenstein, A.H.
J. Biol. Chem. 246, 1365-1374, 1971
A:Title: isolation and characterization of proinsulin C-peptide from bovine pancreas.

A:Reference number: A92074; PMID:7116409; PMID:5545080
A:Accession: A92074
A:Molecule type: protein
A:Residues: 57-82 <STE>
R:Salokangas, A.; Smyth, D.G.; Markussen, J.; Sundby, P.
Eur. J. Biochem. 20, 183-189, 1971
A:Title: Bovine proinsulin: amino acid sequence of the C-peptide isolated from pancreas.
A:Reference number: A91185; PMID:71257721; PMID:5105368
A:Accession: A91185
A:Molecule type: protein
A:Residues: 57-82 <SAL>
R:Sanger, F.; Thompson, E.O.P.
Biochem. J. 53, 366-374, 1953
A:Title: The amino-acid sequence in the glycy chain of insulin. 2. The investigation of
A:Reference number: A90342
A:Accession: A90342
A:Molecule type: protein
A:Residues: 85-105 <SAN>
R:Sanger, F.; Tuppy, H.
Biochem. J. 49, 481-490, 1951
A:Title: The amino-acid sequence in the phenylalanyl chain of insulin. 2. The investigat
A:Reference number: A90341
A:Accession: A90341
A:Molecule type: protein
A:Residues: 25-54 <SA2>
R:Cheng, R.; Kawakishi, S.
Eur. J. Biochem. 223, 759-764, 1994
A:Title: Site-specific oxidation of histidine residues in glycated insulin mediated by
A:Reference number: S48184; PMID:9433378; PMID:8055551
A:Accession: S48184
A:Molecule type: protein
A:Residues: 85-105 <CHE>
A:Accession: S48185
A:Status: preliminary
A:Molecule type: protein
A:Residues: 25-30, 'X', 32-42, 'X', 44-54 <CH2>
R:Lytle, A.P.; Sanger, F.; Smith, L.F.; Kitai, R.
Biochem. J. 60, 541-556, 1955
A:Title: The disulphide bonds of insulin.
A:Reference number: A90343
A:Contents: annotation; amides; disulfides
R:Wentzel, T.; Bockerskorn, C.; Lottspeich, F.; Baumeister, W.
FEBS Lett. 349, 205-209, 1994
A:Title: Existence of a molecular ruler in proteasomes suggested by analysis of degraded
A:Reference number: S46258; PMID:94326921; PMID:8050567
A:Accession: S46258
A:Status: preliminary
A:Molecule type: protein
A:Residues: 25-54 <WEN>
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-54/Domain: insulin chain B #status experimental <BOH>
F:55-84,85-105/Domain: insulin #status experimental <MAT>
F:57-82/Domain: connecting peptide #status experimental <CHP>
F:85-105/Domain: insulin chain A #status experimental <ACH>
F:31-91,43-104,90-95/Disulfide bonds: #status experimental

Query Match 79.2%; Score 366.5; DB 1; Length 105;
Best Local Similarity 80.2%; Pred. No. 1.7e-32;
Matches 69; Conservative 2; Mismatches 10; Indels 5; Gaps 1;

QY 1 FVNQHLGSHLYALTYVCGERGFYTPKRRRAEDLYQVGLGGPGAGSLQPLALEG 60
DB 25 FVNQHLGSHLYALTYVCGERGFYTPKRRRAEDLYQVGLGGPGAGSLQPLALEG 79

QY 61 SLOKRGIVQCCISGLYOLENYCN 86
DB 80 PPOKRGIVQCCISGLYOLENYCN 105

RESULT 15
INMSI


```

insulin 1 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 24-Apr-1984 #sequence_revision 14-Jul-1994 #text_change 09-Jul-2004
C:Accession: B26342; A48172; A01592; B61012
R:Wentworth, B.W.; Schaefer, I.M.; Villa-Komaroff, L.; Chirgwin, J.M.
J. Mol. Evol. 23, 305-312, 1986
A:Title: Characterization of the two nonallelic genes encoding mouse preproinsulin.
A:Reference number: A92965; MUID:87169768; PMID:3104603
A:Accession: B26342
A:Molecule type: DNA
A:Residues: 1-108 <MEN>
A:Cross-references: UNIPROT:P01325; GB:X04725; NID:952712; PIDN:CAA28434.1; PID:952713
R:Sawa, T.; Ohgaku, S.; Moritoka, H.; Yano, S.
J. Mol. Endocrinol. 5, 61-67, 1990
A:Title: Molecular cloning and DNA sequence analysis of preproinsulin genes in the NON
A:Reference number: A48172; MUID:90372989; PMID:2397023
A:Accession: A48172
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-108 <SAW>
R:Buenzli, H.F.; Glatthaar, B.; Kunz, P.; Wuelhaupt, E.; Humbel, R.E.
Hoppe-Seyler's Z. Physiol. Chem. 353, 451-458, 1972
A:Title: Amino acid sequence of the two insulins from mouse (Mus musculus).
A:Reference number: A01592; MUID:72189455; PMID:5063718
A:Accession: A01592
A:Molecule type: protein
A:Residues: 25-54;88-108 <BUS>
R:Linde, S.; Nielsen, J.H.; Hansen, B.; Welinder, B.S.
J. Chromatogr. 462, 243-254, 1989
A:Title: Reversed-phase high-performance liquid chromatographic analyses of insulin bio
A:Reference number: A61012; MUID:89292078; PMID:2661585
A:Accession: B61012
A:Molecule type: protein
A:Residues: 57-85 <LIN>
A:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-26/Domain: signal sequence #status predicted <SIG>
F:25-84/Domain: insulin chain B #status experimental <BGH>
F:25-84,88-108/Product: insulin #status experimental <MAT>
F:57-85/Domain: connecting peptide #status experimental <CPBP>
F:88-108/Domain: insulin chain A #status experimental <ACH>
F:31-94,43-107,93-98/Disulfide bonds: #status predicted

Query Match          79.0%  Score 366; DB 1; Length 108;
Best Local Similarity 81.4%  Pred. No. 2e-32;
Matches 70; Conservative 4; Mismatches 10; Indels 2; Gaps 1;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYPTKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
DB 25 FVKQHLGCPHLVEALYLVCGERGFFYPTKSRREVEDPOVQLTLGSP--GDLQTLALEV 82
QY 61 SLQKRGIVQCCCTSLCSLYQENYCN 86
DB 83 ARQKRGIVDCCCTSLCSLYQENYCN 108

Search completed: November 2, 2004, 20:22:15
Job time : 15.8672 secs

```

[Redacted header line]

[Redacted line of text]

[Redacted body text]

Sat Nov 6 18:59:24 2004

us-10-054-873-4.rapb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 2, 2004, 20:20:47 ; Search time 64.4207 Seconds
(without alignments)
432.820 Million cell updates/sec

Title: US-10-054-873-4
Perfect score: 463
Sequence: 1 FVNHGCGSHLVEALYLVCGRGFFYPKRRAREDDLVQVQLGGPGAGSLQPLALEG 86

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1370721 seqs, 324215800 residues
Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	463	100.0	86	US-09-878-380-1	Sequence 1, Appli
2	463	100.0	86	US-09-858-9358-4	Sequence 4, Appli
3	463	100.0	86	US-10-028-410-2	Sequence 2, Appli
4	463	100.0	86	US-10-054-873-4	Sequence 2, Appli
5	463	100.0	86	US-10-444-326-2	Sequence 2, Appli
6	463	100.0	86	US-10-271-869-4	Sequence 4, Appli
7	463	100.0	86	US-10-444-262-2	Sequence 2, Appli
8	463	100.0	86	US-10-444-649-2	Sequence 2, Appli
9	463	100.0	86	US-10-444-701-2	Sequence 2, Appli
10	463	100.0	96	US-09-947-563-4	Sequence 4, Appli
11	463	100.0	110	US-09-205-658-125	Sequence 125, App
12	463	100.0	110	US-09-815-229-3	Sequence 3, Appli
13	463	100.0	110	US-09-804-409A-9	Sequence 9, Appli

14	463	100.0	110	US-09-969-748C-6	Sequence 6, Appli
15	463	100.0	110	US-09-963-693-125	Sequence 125, App
16	463	100.0	110	US-10-038-686-1	Sequence 1, Appli
17	463	100.0	110	US-10-328-813-2	Sequence 2, Appli
18	463	100.0	110	US-10-383-285-2	Sequence 2, Appli
19	463	100.0	110	US-10-346-563-2	Sequence 2, Appli
20	463	100.0	110	US-10-321-717-2	Sequence 2, Appli
21	463	100.0	110	US-10-411-037-44	Sequence 4, Appli
22	463	100.0	110	US-10-411-026-44	Sequence 4, Appli
23	463	100.0	110	US-10-410-962-44	Sequence 4, Appli
24	463	100.0	110	US-10-411-049-44	Sequence 4, Appli
25	463	100.0	110	US-10-700-725-20	Sequence 20, Appli
26	463	100.0	110	US-10-410-930-44	Sequence 4, Appli
27	463	100.0	110	US-10-410-997-44	Sequence 4, Appli
28	463	100.0	110	US-10-411-012-44	Sequence 4, Appli
29	463	100.0	110	US-10-287-994-44	Sequence 4, Appli
30	463	100.0	110	US-10-740-098-3	Sequence 3, Appli
31	463	100.0	110	US-10-410-913-44	Sequence 4, Appli
32	463	100.0	117	US-08-280-030-63	Sequence 63, Appli
33	463	100.0	130	US-08-280-030-62	Sequence 62, Appli
34	457	98.7	96	US-09-947-563-5	Sequence 5, Appli
35	442	95.5	110	US-10-419-539-5	Sequence 5, Appli
36	438.5	94.7	124	US-10-221-677-24	Sequence 24, Appli
37	306	66.1	166	US-09-925-297-805	Sequence 805, App
38	300	64.8	56	US-09-815-229-10	Sequence 10, Appli
39	300	64.8	56	US-10-740-098-10	Sequence 10, Appli
40	285	61.6	54	US-09-815-229-13	Sequence 13, Appli
41	285	61.6	54	US-10-740-098-13	Sequence 13, Appli
42	267	57.7	52	US-10-054-873-5	Sequence 5, Appli
43	267	57.7	107	US-10-054-873-6	Sequence 6, Appli
44	267	57.7	137	US-10-101-454-39	Sequence 39, Appli
45	267	57.7	145	US-10-101-454-45	Sequence 45, Appli

ALIGNMENTS

RESULT 1
US-09-878-380-1
; Sequence 1, Application US/09878380
; Patent No. US20020160435A1
; GENERAL INFORMATION:
; APPLICANT: Fujirebio Inc.
; APPLICANT: KITAJO, Sachiko
; APPLICANT: KURANO, Yoshihiro
; APPLICANT: NAKATSUBO, Kaoru
; APPLICANT: NISHIZONO, Isao
; TITLE OF INVENTION: Immunoassay For Measuring Human C-Peptide and Kit Therefor
; FILE REFERENCE: 0760-0291P
; CURRENT APPLICATION NUMBER: US/09/878,380
; CURRENT FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: JP 2000-174691
; PRIOR FILING DATE: 2000-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-878-380-1
Query Match 100.0%; Score 463; DB 9; Length 86;
Best Local Similarity 100.0%; Pred. No. 1,2e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FVNHGCGSHLVEALYLVCGRGFFYPKRRAREDDLVQVQLGGPGAGSLQPLALEG 60
DB 1 FVNHGCGSHLVEALYLVCGRGFFYPKRRAREDDLVQVQLGGPGAGSLQPLALEG 60
QY 61 SLQKRGIVQCCCTGCSLYQLENYCN 86
DB 61 SLQKRGIVQCCCTGCSLYQLENYCN 86

RESULT 2
US-09-858-935B-4
Sequence 4, Application US/09858935B
Publication No. US20030069177A1
GENERAL INFORMATION:
APPLICANT: Dubaque, Yves
APPLICANT: Filvaroff, Ellen
APPLICANT: Lowman, Henry B.
TITLE OF INVENTION: METHOD FOR TREATING CARTILAGE DISORDERS
FILE REFERENCE: P1794R1
CURRENT APPLICATION NUMBER: US/09/858,935B
CURRENT FILING DATE: 2002-07-02
PRIOR APPLICATION NUMBER: US 60/248,985
PRIOR FILING DATE: 2000-11-15
PRIOR APPLICATION NUMBER: US 60/204,490
PRIOR FILING DATE: 2000-05-16
NUMBER OF SEQ ID NOS: 153
SEQ ID NO 4
LENGTH: 86
TYPE: PRT
ORGANISM: Homo sapiens
US-09-858-935B-4
Query Match 100.0%; Score 463; DB 10; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.2e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
QY 61 SLQKRGIVQCCCTSGISLYOLENYCN 86
DB 61 SLQKRGIVQCCCTSGISLYOLENYCN 86
RESULT 3
US-10-028-410-2
Sequence 2, Application US/10028410
Publication No. US20020160955A1
GENERAL INFORMATION:
APPLICANT: Dubaque, Yves
APPLICANT: Lowman, Henry
TITLE OF INVENTION: PROTEIN VARIANTS
FILE REFERENCE: P1712R1-1
CURRENT APPLICATION NUMBER: US/10/028,410
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: US/09/477,924
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 6
SEQ ID NO 2
LENGTH: 86
TYPE: PRT
ORGANISM: Homo sapiens
US-10-028-410-2
Query Match 100.0%; Score 463; DB 13; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.2e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
QY 61 SLQKRGIVQCCCTSGISLYOLENYCN 86
DB 61 SLQKRGIVQCCCTSGISLYOLENYCN 86
RESULT 4
US-10-054-873-4
Sequence 4, Application US/10054873

Publication No. US20020164712A1
GENERAL INFORMATION:
APPLICANT: Gan, Zhong Ru
TITLE OF INVENTION: Chimeric Protein Containing an Intramolecular Chaperone-like Sequence
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/054,873
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/CN98/00052
FILING DATE: 31-MAR-1998
APPLICATION NUMBER: US 09/423,100
FILING DATE: 11-DEC-2000
ATTORNEY/AGENT INFORMATION:
NAME: Mycroft, Frank J
REGISTRATION NUMBER: 46,946
REFERENCE/DOCKET NUMBER: 020167-000130US
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 86 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-054-873-4
Query Match 100.0%; Score 463; DB 13; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.2e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
QY 61 SLQKRGIVQCCCTSGISLYOLENYCN 86
DB 61 SLQKRGIVQCCCTSGISLYOLENYCN 86
RESULT 5
US-10-444-326-2
Sequence 2, Application US/10444326
Publication No. US20030191065A1
GENERAL INFORMATION:
APPLICANT: Dubaque, Yves
APPLICANT: Lowman, Henry
TITLE OF INVENTION: PROTEIN VARIANTS
FILE REFERENCE: P1712R1
CURRENT APPLICATION NUMBER: US/10/444,326
CURRENT FILING DATE: 2003-05-22
PRIOR APPLICATION NUMBER: US/09/723,866
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: US/09/477,923
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 6
SEQ ID NO 2
LENGTH: 86
TYPE: PRT

Sat Nov 6 18:59:24 2004

us-10-054-873-4.rapb

Page 3

ORGANISM: Homo sapiens
US-10-444-326-2

Query Match 100.0%; Score 463; DB 14; Length 86;
Best Local Similarity 100.0%; Pred. No. 1,2e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLYEALYVCGERGFYTPKTRREARDLVQGVELGGGPGAGSLQPLALEG 60
DB 1 FVNQHLGSHLYEALYVCGERGFYTPKTRREARDLVQGVELGGGPGAGSLQPLALEG 60
QY 61 SLOKRGIVEQCCTSTICSLYQLENYCN 86
DB 61 SLOKRGIVEQCCTSTICSLYQLENYCN 86

RESULT 6
US-10-271-869-4
Sequence 4, Application US/10271869
Publication No. US20030211992A1
GENERAL INFORMATION:
APPLICANT: Dubague, Yves
APPLICANT: Filvaroff, Ellen
APPLICANT: Lowman, Henry B.
TITLE OF INVENTION: METHOD FOR TREATING CARTILAGE DISORDERS
FILE REFERENCE: P1794R1
CURRENT APPLICATION NUMBER: US/10/271,869
CURRENT FILING DATE: 2002-10-16
PRIOR APPLICATION NUMBER: US/09/858,935
PRIOR FILING DATE: 2002-07-02
PRIOR APPLICATION NUMBER: US 60/248,985
PRIOR FILING DATE: 2000-11-15
PRIOR APPLICATION NUMBER: US 60/204,490
PRIOR FILING DATE: 2000-05-16
NUMBER OF SEQ ID NOS: 153
SEQ ID NO 4
LENGTH: 86
TYPE: PRT
ORGANISM: Homo sapiens
US-10-271-869-4

Query Match 100.0%; Score 463; DB 14; Length 86;
Best Local Similarity 100.0%; Pred. No. 1,2e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLYEALYVCGERGFYTPKTRREARDLVQGVELGGGPGAGSLQPLALEG 60
DB 1 FVNQHLGSHLYEALYVCGERGFYTPKTRREARDLVQGVELGGGPGAGSLQPLALEG 60
QY 61 SLOKRGIVEQCCTSTICSLYQLENYCN 86
DB 61 SLOKRGIVEQCCTSTICSLYQLENYCN 86

RESULT 7
US-10-444-262-2
Sequence 2, Application US/10444262
Publication No. US20040023883A1
GENERAL INFORMATION:
APPLICANT: Dubague, Yves
APPLICANT: Lowman, Henry
TITLE OF INVENTION: PROTEIN VARIANTS
FILE REFERENCE: P1712R1
CURRENT APPLICATION NUMBER: US/10/444,262
CURRENT FILING DATE: 2003-05-22
PRIOR APPLICATION NUMBER: US/09/724,478
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: US/09/477,923
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 6
SEQ ID NO 2
LENGTH: 86
TYPE: PRT
ORGANISM: Homo sapiens
US-10-444-262-2

ORGANISM: Homo sapiens
US-10-444-262-2

Query Match 100.0%; Score 463; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 1,2e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLYEALYVCGERGFYTPKTRREARDLVQGVELGGGPGAGSLQPLALEG 60
DB 1 FVNQHLGSHLYEALYVCGERGFYTPKTRREARDLVQGVELGGGPGAGSLQPLALEG 60
QY 61 SLOKRGIVEQCCTSTICSLYQLENYCN 86
DB 61 SLOKRGIVEQCCTSTICSLYQLENYCN 86

RESULT 8
US-10-444-649-2
Sequence 2, Application US/10444649
Publication No. US20040033951A1
GENERAL INFORMATION:
APPLICANT: Dubague, Yves
APPLICANT: Lowman, Henry
TITLE OF INVENTION: PROTEIN VARIANTS
FILE REFERENCE: P1712R1
CURRENT APPLICATION NUMBER: US/10/444,649
CURRENT FILING DATE: 2003-05-22
PRIOR APPLICATION NUMBER: US/09/724,479
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: US/09/477,923
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 6
SEQ ID NO 2
LENGTH: 86
TYPE: PRT
ORGANISM: Homo sapiens
US-10-444-649-2

Query Match 100.0%; Score 463; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 1,2e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLYEALYVCGERGFYTPKTRREARDLVQGVELGGGPGAGSLQPLALEG 60
DB 1 FVNQHLGSHLYEALYVCGERGFYTPKTRREARDLVQGVELGGGPGAGSLQPLALEG 60
QY 61 SLOKRGIVEQCCTSTICSLYQLENYCN 86
DB 61 SLOKRGIVEQCCTSTICSLYQLENYCN 86

RESULT 9
US-10-444-701-2
Sequence 2, Application US/10444701
Publication No. US20040033952A1
GENERAL INFORMATION:
APPLICANT: Dubague, Yves
APPLICANT: Lowman, Henry
TITLE OF INVENTION: PROTEIN VARIANTS
FILE REFERENCE: P1712R1
CURRENT APPLICATION NUMBER: US/10/444,701
CURRENT FILING DATE: 2003-05-22
PRIOR APPLICATION NUMBER: US/09/723,866
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: US/09/477,923
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 6
SEQ ID NO 2
LENGTH: 86
TYPE: PRT
ORGANISM: Homo sapiens
US-10-444-701-2

Query Match 100.0%; Score 463; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 1,2e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRRAEDLQGVQLGGGPGAGSLQPLALEG 60
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRRAEDLQGVQLGGGPGAGSLQPLALEG 60
QY 61 SLQKRGIVEOCCTISCSLYOLENYCN 86
DB 61 SLQKRGIVEOCCTISCSLYOLENYCN 86

RESULT 10

US-09-947-563-4
Sequence 4, Application US/09947563
Patent No. US20020156234A1
GENERAL INFORMATION:

APPLICANT: Rubroder, Franz-Josef

Keller, Reinhold

TITLE OF INVENTION: Improved process for obtaining
insulin precursors having correctly bonded cystine bridges

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegun, Henderson, Farrabow, Garrett &

Dunner

STREET: 1300 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/947,563

FILING DATE: 07-Sep-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/134,836

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Leslie McDonnell

REGISTRATION NUMBER: 34,872

REFERENCE/DOCKET NUMBER: 02481.1600-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 408-4000

TELEFAX: (202) 408-4400

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 96 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: Escherichia coli

FEATURE:

NAME/KEY: Protein

LOCATION: 1..96

SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-947-563-4

Query Match 100.0%; Score 463; DB 9; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.4e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRRAEDLQGVQLGGGPGAGSLQPLALEG 60
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRRAEDLQGVQLGGGPGAGSLQPLALEG 70

QY 61 SLQKRGIVEOCCTISCSLYOLENYCN 86
DB 71 SLQKRGIVEOCCTISCSLYOLENYCN 96

RESULT 11

US-09-205-658-125
Sequence 125, Application US/09205658
Patent No. US20010029617A1
GENERAL INFORMATION:

APPLICANT: Ruvkun, Gary

APPLICANT: Ogg, Scott

TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
IMPAIRED GLUCOSE TOLERANCE CONDITIONS

FILE REFERENCE: 00786/35104

CURRENT APPLICATION NUMBER: US/09/205,658

CURRENT FILING DATE: 1998-12-03

EARLIER APPLICATION NUMBER: 08/857,076

EARLIER FILING DATE: 1997-05-15

EARLIER APPLICATION NUMBER: 08/888,534

EARLIER FILING DATE: 1997-07-07

EARLIER APPLICATION NUMBER: US98/10080

EARLIER FILING DATE: 1998-05-15

NUMBER OF SEQ ID NOS: 328

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO: 125

LENGTH: 110

TYPE: PRT

ORGANISM: Homo sapiens

US-09-205-658-125

Query Match 100.0%; Score 463; DB 9; Length 110;
Best Local Similarity 100.0%; Pred. No. 1.6e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRRAEDLQGVQLGGGPGAGSLQPLALEG 60
DB 25 FVNQHLGSHLVEALYLVCGERGFFYTPKTRRAEDLQGVQLGGGPGAGSLQPLALEG 84
QY 61 SLQKRGIVEOCCTISCSLYOLENYCN 86
DB 85 SLQKRGIVEOCCTISCSLYOLENYCN 110

RESULT 12

US-09-815-229-3
Sequence 3, Application US/09815229
Patent No. US20020058614A1
GENERAL INFORMATION:

APPLICANT: Filvaroff, Ellen H.

APPLICANT: Okumu, Franklin W.

TITLE OF INVENTION: USE OF INSULIN FOR THE TREATMENT OF CARTILAGEOUS DISORDERS

FILE REFERENCE: P1786R1US

CURRENT APPLICATION NUMBER: US/09/815,229

CURRENT FILING DATE: 2001-03-22

PRIOR APPLICATION NUMBER: US 60/192,103

PRIOR FILING DATE: 2000-03-24

NUMBER OF SEQ ID NOS: 17

SEQ ID NO: 3

LENGTH: 110

TYPE: PRT

ORGANISM: Homo sapiens

US-09-815-229-3

Query Match 100.0%; Score 463; DB 9; Length 110;
Best Local Similarity 100.0%; Pred. No. 1.6e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRRAEDLQGVQLGGGPGAGSLQPLALEG 60
DB 25 FVNQHLGSHLVEALYLVCGERGFFYTPKTRRAEDLQGVQLGGGPGAGSLQPLALEG 84
QY 61 SLQKRGIVEOCCTISCSLYOLENYCN 86

Sat Nov 6 18:59:24 2004

ua-10-054-873-4.rapb

Page 5

Db 85 SLQKRGIVECCCTICSISLYQLENYCN 110

RESULT 13
US-09-804-409A-9
Sequence 9, Application US/09804409A
Patent No. US20020155100A1
GENERAL INFORMATION:
APPLICANT: KIEFER, TIMOTHY J.
APPLICANT: CHEUNG, ANTHONY T.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATED PROTEIN
FILE REFERENCE: 029996/027 8721
CURRENT APPLICATION NUMBER: US/09/804,409A
CURRENT FILING DATE: 2001-03-12
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn ver. 2.1
SEQ ID NO 9
LENGTH: 110
TYPE: PRT
ORGANISM: Homo sapiens
US-09-804-409A-9

Query Match 100.0%; Score 463; DB 9; Length 110;
Best Local Similarity 100.0%; Pred. No. 1.6e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQVELGGGPGAGSIQPLALEG 60
Db 25 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQVELGGGPGAGSIQPLALEG 84
Qy 61 SLQKRGIVECCCTICSISLYQLENYCN 86
Db 85 SLQKRGIVECCCTICSISLYQLENYCN 110

RESULT 14
US-09-969-748C-6
Sequence 6, Application US/09969748C
Publication No. US20030161809A1
GENERAL INFORMATION:
APPLICANT: ARIZBEK PHARMACEUTICALS, INC.
APPLICANT: HOUSTON, Lou, L.
APPLICANT: SHERIDAN, Philip, J.
APPLICANT: HAMLEY, Stephen
APPLICANT: GLYNN, Jacqueline, M.
APPLICANT: CHAPIN, Steven
APPLICANT: BASU, Amareesh
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TRANSPORT OF BIOLOGICALLY ACTIVE
FILE REFERENCE: 057220-0303
CURRENT APPLICATION NUMBER: US/09/969,748C
CURRENT FILING DATE: 2002-12-10
PRIOR APPLICATION NUMBER: US 60/267,601
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: US 60/248,819
PRIOR FILING DATE: 2000-11-14
PRIOR APPLICATION NUMBER: US 60/248,478
PRIOR FILING DATE: 2000-11-13
PRIOR APPLICATION NUMBER: US 60/237,929
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 115
SOFTWARE: PatentIn version 3.0
SEQ ID NO 6
LENGTH: 110
TYPE: PRT
ORGANISM: Homo sapiens
US-09-969-748C-6

Query Match 100.0%; Score 463; DB 10; Length 110;
Best Local Similarity 100.0%; Pred. No. 1.6e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQVELGGGPGAGSIQPLALEG 60
Db 25 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQVELGGGPGAGSIQPLALEG 84
Qy 61 SLQKRGIVECCCTICSISLYQLENYCN 86
Db 85 SLQKRGIVECCCTICSISLYQLENYCN 110

RESULT 15
US-09-963-693-125
Sequence 125, Application US/09963693
Publication No. US20030181364A1
GENERAL INFORMATION:
APPLICANT: Ruwkun, Gary
APPLICANT: Ogas, Scott
TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
FILE REFERENCE: 00786/351004
CURRENT APPLICATION NUMBER: US/09/963,693
CURRENT FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US/09/205,658
PRIOR FILING DATE: 1998-12-03
PRIOR APPLICATION NUMBER: 08/857,076
PRIOR FILING DATE: 1997-05-15
PRIOR APPLICATION NUMBER: 08/888,534
PRIOR FILING DATE: 1997-07-07
PRIOR APPLICATION NUMBER: US96/10080
PRIOR FILING DATE: 1996-05-15
NUMBER OF SEQ ID NOS: 328
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 125
LENGTH: 110
TYPE: PRT
ORGANISM: Homo sapiens
US-09-963-693-125

Query Match 100.0%; Score 463; DB 10; Length 110;
Best Local Similarity 100.0%; Pred. No. 1.6e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQVELGGGPGAGSIQPLALEG 60
Db 25 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQVELGGGPGAGSIQPLALEG 84
Qy 61 SLQKRGIVECCCTICSISLYQLENYCN 86
Db 85 SLQKRGIVECCCTICSISLYQLENYCN 110

Search completed: November 2, 2004, 20:59:22
Job time : 65.4207 secs

4

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 2, 2004, 19:48:36 ; Search time 83.3026 Seconds
(without alignments)
594.006 Million cell updates/sec

Title: US-10-054-873-4

Perfect score: 463
Sequence: 1 FVNHQICGSHLWALYIVCG.....IVEQCCTSLCYLQENYCN 86

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: Uniprot_02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	463	100.0	110 1	INS_HUMAN P01308 homo sapien
2	463	100.0	110 1	INS_PANTR P01410 pan troglod
3	463	100.0	110 1	INS_PONPY Q6hxy2 pongo pygma
4	463	100.0	110 2	Q6yK33 gorilla gor
5	463	100.0	110 2	AAP35454 Aap35454 homo sapi
6	463	100.0	110 2	AAN06935 AAN06935 gorilla g
7	463	100.0	110 2	AAN39451 AAN39451 homo sapi
8	456	98.5	110 1	INS_CERAB P30407 cercopithe
9	456	98.5	110 1	INS_MACPA P30406 macaca fasc
10	424	91.6	110 1	INS_RABIT P01311 coryctolagu
11	417	90.1	110 1	INS_CANPA P01321 canis fami
12	413	89.2	110 1	INS_SPERT P01313 spermophilu
13	394	85.1	110 1	INS_HORSE P01310 equus caball
14	394	85.1	110 1	INS2_MOUSE P01326 mus musculu
15	394	85.1	110 2	INS2_RAT P01323 rattus norv
16	394	85.1	110 2	BAB25135 Bab25135 mus muscu
17	392	84.7	108 1	INS_ACTR P10604 actus trivi
18	392	84.7	110 1	INS_CRITIO P01313 cricetulus
19	388	83.8	110 2	Q8H286 Q8H286 felis silve
20	388	83.2	110 2	P01322 rattus norv
21	383	82.7	108 1	INS_FIG P01315 sus scrofa
22	383	82.7	108 2	AAQ00952 AAQ00952 sus scrof
23	383	82.7	108 2	AAQ00954 AAQ00954 sus scrof
24	383	82.7	108 2	AAQ00957 AAQ00957 sus scrof
25	383	82.7	108 2	AAQ00960 AAQ00960 sus scrof
26	383	82.7	108 2	AAQ00963 AAQ00963 sus scrof
27	383	82.7	108 2	AAQ00966 AAQ00966 sus scrof
28	383	82.7	108 2	AAQ00969 AAQ00969 sus scrof
29	383	82.7	108 2	AAQ00972 AAQ00972 sus scrof
30	383	82.7	108 2	AAQ00975 AAQ00975 sus scrof
31	383	82.7	108 2	AAQ00978 AAQ00978 sus scrof

32	383	82.7	108 2	AAQ00981 sus scrof
33	383	82.7	108 2	AAQ00983 sus scrof
34	383	82.7	108 2	AAQ00985 sus scrof
35	383	82.7	108 2	AAQ00987 sus scrof
36	383	82.7	108 2	AAQ00990 sus scrof
37	377	81.4	110 1	INS_PSAOB Q62587 psammomy o
38	366.5	79.2	105 1	INS_BOVIN P01317 bos taurus
39	366	79.0	108 1	INS_MOUSE P01325 mus musculu
40	366	79.0	108 2	BAB24974 Bab24974 mus muscu
41	366	79.0	108 2	BAB25628 mus muscu
42	362.5	78.3	105 1	INS_SHEEP P01318 ovis aries
43	342	73.9	65 2	Q8H280 Q8H280 pongo pygma
44	342	73.9	65 2	Q8H281 Q8H281 gorilla gor
45	334.5	72.2	108 1	INS_RODSP P15163 rodentia sp

ALIGNMENTS

RESULT 1
INS_HUMAN STANDARD; PRT; 110 AA.
ID INS_HUMAN
AC P01308;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Insulin precursor.
GN Name-INS;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=60243748; PubMed=6243748;
RA Bell G.I., Picet R.L., Rutter W.J., Cordell B., Tischer E.,
RA Goodman H.M., "Sequence of the human insulin gene.",
RL Nature 284:26-32(1980).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=6023633; PubMed=6248962;
RA Ulrich A., Dull T.J., Gray A., Brosius J., Sures I.,
RL "Genetic variation in the human insulin gene.",
RL Science 209:612-615(1980).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=60054779; PubMed=503234;
RA Bell G.I., Swain W.F., Picet R.L., Cordell B., Goodman H.M.,
RA Rutter W.J.,
RL "Nucleotide sequence of a cDNA clone encoding human preproinsulin.",
RL Nature 282:525-527(1979).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=60147417; PubMed=6927840;
RA Sures I., Goeddel D.V., Gray A., Ulrich A.,
RL "Nucleotide sequence of human preproinsulin complementary DNA.",
RL Science 208:57-59(1980).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=63364428; PubMed=8358440;
RA Innesen A.M., Bell J.I., Jullier C., Lathrop M.,
RL "Susceptibility to insulin dependent diabetes mellitus maps to a 4.1
kb segment of DNA spanning the insulin gene and associated VNTR.",
RL Nat. Genet. 4:305-310(1993).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shamen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Wang J., Hsieh F.,

- RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo K.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinot P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Besak S.A., Mcwan P.J., McKernan K.J., Malek J.A., Guneratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Murry D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schmeckel A., Schein J.E., Jones S.J.M., Marra M.A.,
RA "Generation and initial analysis of more than 15,000 full-length human
RA and mouse cDNA sequences.";
RA Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RA [17]
RA SEQUENCE OF 1-59 FROM N.A.
RA TISSUE=Blood;
RA RefSeq=1.1; Weill J.J., Stuckens C.C., Danze P.M.P.,
RA "Description of a novel RFLP diallelic polymorphism (-127 BsgI C/G)
RA within the 5' region of insulin gene.";
RA Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RA [18]
RA SEQUENCE OF 25-54 AND 90-110.
RA Nicol D.S.H.W., Smith L.F.,
RA "Amino-acid sequence of human insulin.";
RA Nature 187:483-485(1960).
RA [19]
RA SEQUENCE OF 57-87.
RA MEDLINE=7116410; PubMed=5101771;
RA Oyer P.E., Cho S., Peterson J.D., Steiner D.F.,
RA "Studies on human proinsulin. Isolation and amino acid sequence of the
RA human pancreatic C-peptide.";
RA J Biol. Chem. 246:1375-1386(1971).
RA [10]
RA SEQUENCE OF 57-87.
RA MEDLINE=7125722; PubMed=5560404;
RA Ko A., Smyth D.G., Markussen J., Sundby F.,
RA "The amino acid sequence of the C-peptide of human proinsulin.";
RA Eur. J. Biochem. 20:190-199(1971).
RA [11]
RA SEQUENCE OF 57-87.
RA MEDLINE=7507277; PubMed=4443293;
RA Steber P., Kamber B., Hartmann A., Joehl A., Riniker B., Rittler W.,
RA "Total synthesis of human insulin under directed formation of the
RA disulfide bonds.";
RA Helv. Chim. Acta 57:2617-2621(1974).
RA [12]
RA SEQUENCE OF 57-87.
RA MEDLINE=75040007; PubMed=4803504;
RA Nathan V.K.,
RA "Studies on polypeptides. IV. The synthesis of C-peptide of human
RA proinsulin.";
RA Hoppe-Seyler's Z. Physiol. Chem. 354:659-672(1973).
RA [13]
RA SEQUENCE OF 65-69 AND 70-73.
RA MEDLINE=7316263; PubMed=4698555;
RA Geiger R., Volk A.,
RA "Synthesis of peptides with the properties of human proinsulin C
RA peptides (hc peptide). 3 Synthesis of the sequences 14-17 and 9-13 of
RA human proinsulin C peptides.";
RA Chem. Ber. 106:199-205(1973).
RA [14]
RA SEQUENCE OF 84-87.
RA MEDLINE=7316261; PubMed=4698553;
RA Geiger R., Jaeger G., Keonig W., Treuth G.,
RA "Synthesis of peptides with the properties of human proinsulin C
RA peptides (hc peptide). I. Scheme for the synthesis and preparation of
RA the sequence 26-31 of human proinsulin C peptide.";
RA Chem. Ber. 106:188-192(1973).
RA [15]
RA VARIANT LOS ANGELES SER-48.
RX MEDLINE=84016053; PubMed=6312455;
RA Haneida M., Chan S.J., Kwok S.C.M., Rubenstein A.H., Steiner D.F.,
RT "Studies on mutant human insulin genes: identification and sequence
RT analysis of a gene encoding [Ser24]insulin.";
RT Proc. Natl. Acad. Sci. U.S.A. 80:6366-6370(1983).
RX [16]
RX MEDLINE=84170233; PubMed=6424111;
RA Shelson S., Fickova M., Haneida M., Nahum A., Musso G., Kaiser E.T.,
RA Rubenstein A.H., Tager H.,
RT "Identification of a mutant human insulin predicted to contain a
RT serine-for-phenylalanine substitution.";
RT Proc. Natl. Acad. Sci. U.S.A. 80:7390-7394(1983).
RX [17]
RX MEDLINE=87175640; PubMed=3470784;
Chan S.J., Seino S., Grunpuro P.A., Schwartz R., Steiner D.F.,
RT "A mutation in the B chain coding region is associated with impaired
RT proinsulin conversion in a family with hyperproinsulinemia.";
RT Proc. Natl. Acad. Sci. U.S.A. 84:2194-2197(1987).
RX [18]
RX MEDLINE=87058122; PubMed=3537011;
Sakura H., Iwamoto Y., Sakamoto Y., Kuzuya T., Hirata H.,
RT "Structurally abnormal insulin in a diabetic patient. Characterization
RT of the mutant insulin A3 (Val--2Ieu) isolated from the pancreas.";
RT J Clin. Invest. 78:1666-1672(1986).
RX [19]
RX MEDLINE=90317021; PubMed=2196279;
Barbetti F., Raben N., Kadowaki T., Cama A., Accili D., Gabbay K.H.,
RA Mereshki J.A., Taylor S.I., Roth J.,
RT "Two unrelated patients with familial hyperproinsulinemia due to a
RT mutation substituting histidine for arginine at position 65 in the
RT proinsulin molecule: identification of the mutation by direct
RT sequencing of genomic deoxyribonucleic acid amplified by polymerase
RT chain reaction.";
RT J. Clin. Endocrinol. Metab. 71:164-169(1990).
RX [20]
RX MEDLINE=8526196; PubMed=4019786;
Shibasaki Y., Kawakami T., Kanazawa Y., Akanuma Y., Takaku F.,
RT "Posttranslational cleavage of proinsulin is blocked by a point
RT mutation in familial hyperproinsulinemia.";
RT J Clin. Invest. 76:378-380(1985).
RX [21]
RX MEDLINE=92291307; PubMed=1601997;
Yano H., Kitano N., Morimoto M., Polonsky K.S., Imura H., Seino Y.,
RT "A novel point mutation in the human insulin gene giving rise to
RT hyperproinsulinemia [proinsulin Kyoto].";
RT J Clin. Invest. 89:1902-1907(1992).
RX [22]
RX MEDLINE=91104966; PubMed=2271664;
Hua Q.-X., Weiss M.A.,
RT "Toward the solution structure of human insulin: sequential 2D 1H NMR
RT assignment of a des-pentapeptide analogue and comparison with crystal
RT structure.";
RT Biochemistry 29:10545-10555(1990).
RX [23]
RX MEDLINE=9124467; PubMed=2036420;
Hua Q.-X., Weiss M.A.,
RT "Comparative 2D NMR studies of human insulin and des-pentapeptide
RT insulin: sequential resonance assignment and implications for protein
RT dynamics and receptor recognition.";
RT Biochemistry 30:5505-5515(1991).
RX [24]
RX MEDLINE=91265527; PubMed=1646635;
Hua Q.-X., Weiss M.A.,
RT "Two-dimensional NMR studies of Des-(B26-B30)-insulin: sequence-

RT specific resonance assignments and effects of solvent composition.",
 RL Biochim. Biophys. Acta 1078:101-110(1991).

Query Match 100.0%; Score 463; DB 1; Length 110;
 Best Local Similarity 100.0%; Pred. No. 7.2e-41;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEALYVGEGRFYPTKTRAEADLVQGVLEGGPGAGSLQPLALEG 60
 DB 25 FVNHLCGSHLVEALYVGEGRFYPTKTRAEADLVQGVLEGGPGAGSLQPLALEG 84
 QY 61 SLOKRGIVEOCCTSCISLYOLENYCN 86
 DB 85 SLOKRGIVEOCCTSCISLYOLENYCN 110

RESULT 2
 INS_PANTR STANDARD; PRT; 110 AA.

AC P30410;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Insulin precursor.
 GN Name=INS;
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
 OX NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92219553; PubMed=1560757;
 RA Seldin S., Bell G.I., Li W.;
 RT "Sequences of primate insulin genes support the hypothesis of a slower
 RT rate of molecular evolution in humans and apes than in monkeys.",
 RL Mol. Biol. Evol. 9:193-203(1992).
 [2]

RP SEQUENCE FROM N.A.
 RX MEDLINE=26833521; PubMed=12952878; DOI=10.1101/gr.948003;
 RA Stead J.D.H., Hurles M.E., Jeffreys A.J.;
 RT "Global haplotype diversity in the human insulin gene region.",
 RL Genome Res. 13:2101-2111(2003).
 CC -1- FUNCTION: Insulin decreases blood glucose concentration. It
 CC increases cell permeability to monosaccharides, amino acids and
 CC fatty acids. It accelerates glycolysis, the pentose phosphate
 CC cycle, and glycogen synthesis in liver.
 CC -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
 CC disulfide bonds.
 CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- SIMILARITY: Belongs to the insulin family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL, X61089; CAA43403.1; -;
 DR EMBL, AY137497; AAN06933.1; -;
 DR PIR, A42179; A42179.
 DR HSSP, P01308; IAI0.
 DR InterPro: IPR004825; Ins/IGF/relax.
 DR Pfam: PF00049; Insulin; 1.
 DR PRINTS, PR00277; INSULINB.
 DR PROSITE, PS00262; INSULIN; 1.
 KW Glucose metabolism; Hormone; Insulin family; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 54 Insulin B chain.
 FT PROPEP 57 87 C peptide.
 FT CHAIN 90 110 Insulin A chain.
 FT DISULFID 31 96 Interchain.

FT DISULFID 43 109 Interchain.

FT DISULFID 95 100
 FT HELIX 32 43
 FT TURN 60 61
 FT TURN 60 61
 FT HELIX 71 71
 FT HELIX 76 82
 SQ SEQUENCE 110 AA; 12025 MW; 41B8DF79837CE5 CRC64;

Query Match 100.0%; Score 463; DB 1; Length 110;
 Best Local Similarity 100.0%; Pred. No. 7.2e-41;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEALYVGEGRFYPTKTRAEADLVQGVLEGGPGAGSLQPLALEG 60
 DB 25 FVNHLCGSHLVEALYVGEGRFYPTKTRAEADLVQGVLEGGPGAGSLQPLALEG 84
 QY 61 SLOKRGIVEOCCTSCISLYOLENYCN 86
 DB 85 SLOKRGIVEOCCTSCISLYOLENYCN 110

RESULT 3
 INS_PONPY STANDARD; PRT; 110 AA.

AC Q8HXV2;
 DT 05-JUL-2004 (Rel. 44, Created)
 DT 05-JUL-2004 (Rel. 44, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Insulin precursor.
 GN Name=INS;
 OS Pongo pygmaeus (Orangutan).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pongo.
 OX NCBI_TaxID=9600;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=26833521; PubMed=12952878; DOI=10.1101/gr.948003;
 RA Stead J.D.H., Hurles M.E., Jeffreys A.J.;
 RT "Global haplotype diversity in the human insulin gene region.",
 RL Genome Res. 13:2101-2111(2003).

CC -1- FUNCTION: Insulin decreases blood glucose concentration. It
 CC increases cell permeability to monosaccharides, amino acids and
 CC fatty acids. It accelerates glycolysis, the pentose phosphate
 CC cycle, and glycogen synthesis in liver.
 CC -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
 CC disulfide bonds.
 CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- SIMILARITY: Belongs to the insulin family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL, AY137503; AAN06937.1; -;
 DR HSSP, P01308; IAI0.
 DR InterPro: IPR004825; Ins/IGF/relax.
 DR Pfam: PF00049; Insulin; 1.
 DR PRINTS, PR00277; INSULINB.
 DR SMART, SM00078; IIGF; 1.
 DR PROSITE, PS00262; INSULIN; 1.
 KW Glucose metabolism; Hormone; Insulin family; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 54 Insulin B chain.
 FT PROPEP 57 87 C peptide.
 FT CHAIN 90 110 Insulin A chain.
 FT DISULFID 31 96 Interchain (By similarity).
 FT DISULFID 43 109 Interchain (By similarity).
 FT DISULFID 95 100 By similarity.
 SQ SEQUENCE 110 AA; 12038 MW; 22D2B32B94F520F8 CRC64;

Query Match 100.0%; Score 463; DB 1; Length 110;
 Best Local Similarity 100.0%; Pred. No. 7.2e-41;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALIVCGERGFYTPKTRREARDLVQGVLEGGPGAGSLQPLALEG 60
 DB 25 FVNQHLGSHLVEALIVCGERGFYTPKTRREARDLVQGVLEGGPGAGSLQPLALEG 84

QY 61 SLQKRGIVGECCTISCSLYQLENYCN 86
 DB 85 SLQKRGIVGECCTISCSLYQLENYCN 110

RESULT 4
 ID 06YK33 PRELIMINARY; PRT; 110 AA.
 AC 06YK33;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE Insulin.
 GN Name=INS;
 OS Gorilla gorilla (gorilla).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
 OC NCBI_TaxID=9593;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22833521; PubMed=12952878;
 RA Stead J.D., Hurles M.E., Jeffreys A.J.;
 RT "Global haplotype diversity in the human insulin gene region."
 RL Genome Res. 13:2101-2111(2003).
 DE Insulin.
 GN [2]
 RP SEQUENCE FROM N.A.
 RA Stead J.D.H., Jeffreys A.J.;
 RC Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- SIMILARITY: Belongs to the insulin family.
 DR EMBL: AY137500; AA066935.1; -;
 DR InterPro: IPR004825; Ins/IGF/relax.
 DR InterPro: IPR003234; Molusc_ins.
 DR Pfam: PF00649; Insulin; 1.
 DR PRINTS: PR00277; INSULIN.
 DR PRODOM: PD015667; Molusc_ins; 1.
 DR SMART: SM00078; IIGF; 1.
 DR PROSITE: PS00262; INSULIN; 1.
 KM Insulin family.
 SQ SEQUENCE 110 AA; 11981 MW; C2C3B23B85E520E5 CRC64;

Query Match 100.0%; Score 463; DB 2; Length 110;
 Best Local Similarity 100.0%; Pred. No. 7.2e-41;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALIVCGERGFYTPKTRREARDLVQGVLEGGPGAGSLQPLALEG 60
 DB 25 FVNQHLGSHLVEALIVCGERGFYTPKTRREARDLVQGVLEGGPGAGSLQPLALEG 84

QY 61 SLQKRGIVGECCTISCSLYQLENYCN 86
 DB 85 SLQKRGIVGECCTISCSLYQLENYCN 110

RESULT 5
 ID AAP35454 PRELIMINARY; PRT; 110 AA.
 AC AAP35454;
 DT 02-MAR-2004 (TREMBlrel. 27, Created)
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
 DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
 DE Insulin.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9605;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kaidima N., Chen X., Rolfe A., Halleck A., Hines L., Eisenstein S.,
 RA Koundinya W., Raphael J., Moreira D., Kelley T., Labaer J., Lin Y.,
 RA Pheasant M., Farmer A.;
 RT "Cloning of human full-length cDNAs in BD Creator (TM) System Donor
 vector."
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BT006808; AAP35454.1; -;
 SQ SEQUENCE 110 AA; 11981 MW; C2C3B23B85E520E5 CRC64;

Query Match 100.0%; Score 463; DB 2; Length 110;
 Best Local Similarity 100.0%; Pred. No. 7.2e-41;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALIVCGERGFYTPKTRREARDLVQGVLEGGPGAGSLQPLALEG 60
 DB 25 FVNQHLGSHLVEALIVCGERGFYTPKTRREARDLVQGVLEGGPGAGSLQPLALEG 84

QY 61 SLQKRGIVGECCTISCSLYQLENYCN 86
 DB 85 SLQKRGIVGECCTISCSLYQLENYCN 110

RESULT 6
 ID AAN06935 PRELIMINARY; PRT; 110 AA.
 AC AAN06935;
 DT 02-MAR-2004 (TREMBlrel. 27, Created)
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
 DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
 DE Insulin.
 GN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22833521; PubMed=12952878;
 RA Stead J.D., Hurles M.E., Jeffreys A.J.;
 RT "Global haplotype diversity in the human insulin gene region."
 RL Genome Res. 13:2101-2111(2003).
 DE Insulin.
 GN [2]
 RP SEQUENCE FROM N.A.
 RA Stead J.D.H., Jeffreys A.J.;
 RC Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY137500; AAN06935.1; -;
 SQ SEQUENCE 110 AA; 11981 MW; C2C3B23B85E520E5 CRC64;

Query Match 100.0%; Score 463; DB 2; Length 110;
 Best Local Similarity 100.0%; Pred. No. 7.2e-41;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALIVCGERGFYTPKTRREARDLVQGVLEGGPGAGSLQPLALEG 60
 DB 25 FVNQHLGSHLVEALIVCGERGFYTPKTRREARDLVQGVLEGGPGAGSLQPLALEG 84

QY 61 SLQKRGIVGECCTISCSLYQLENYCN 86
 DB 85 SLQKRGIVGECCTISCSLYQLENYCN 110

RESULT 7
 ID AAN39451 PRELIMINARY; PRT; 110 AA.
 AC AAN39451;
 DT 02-MAR-2004 (TREMBlrel. 27, Created)
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
 DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
 DE Insulin.

```

CN      INS.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primata; Catarrhini; Hominidae; Homo.
CX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=22833521; PubMed=12952878;
RA      Stead J.D., Hurles M.E., Jeffreys A.J.;
RT      "Global haplotype diversity in the human insulin gene region.";
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Stead J.D.H., Jeffreys A.J.;
RT      Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AY138590; AAN39451.1;
SQ      SEQUENCE 110 AA; 11981 MW; C2C3B23B85E52D5E5 CRC64;

Query March 100.0%; Score 463; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 7,2e-41;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy      1 FVNQHLGSHLVEALYLVCGERGFFYTPKTEREADELQVQVELGGGPGAGSIQAPALEG 60
Db      25 FVNQHLGSHLVEALYLVCGERGFFYTPKTEREADELQVQVELGGGPGAGSIQAPALEG 84
Qy      61 SLQKRGIVGOCCTISLYQENYCN 86
Db      85 SLQKRGIVGOCCTISLYQENYCN 110

RESULT 8
INS_CERAE
ID      INS_CERAE          STANDARD;          PRT;          110 AA.
AC      P30407; P01309;
DT      01-APR-1993 (Rel. 25, Created)
DT      01-APR-1993 (Rel. 25, Last sequence update)
DT      03-JUL-2004 (Rel. 44, Last annotation update)
DE      Insulin precursor.
OS      Name=INS;
OC      Cercopithecus aethiops (Green monkey) (Grivet).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC      Cercopithecinae; Cercopithecus.
OX      NCBI_TaxID=9534;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=9221953; PubMed=1560757;
RA      Saino S., Bell G.I., Li W.;
RT      "Sequences of primate insulin genes support the hypothesis of a slower
RT      rate of molecular evolution in humans and apes than in monkeys.";
RN      [2]
RP      Mol. Biol. Evol. 9:193-203(1992).

RP      SEQUENCE OF 57-87.
RX      MEDLINE=72258016; PubMed=4626369;
RA      Peterson J.D., Nehrlich S., Oyer P.E., Steiner D.F.;
RT      "Determination of the amino acid sequence of the monkey, sheep, and
RT      dog proinsulin C-peptides by a semi-micro Edman degradation
RT      procedure.";
RN      J. Biol. Chem. 247:4666-4871(1972).
CC      -1- FUNCTION: Insulin decreases blood glucose concentration. It
CC      increases cell permeability to monosaccharides, amino acids and
CC      fatty acids. It accelerates glycolysis, the pentose phosphate
CC      cycle, and glycogen synthesis in liver.
CC      -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
CC      disulfide bonds.
CC      -1- SUBCELLULAR LOCATION: Secreted.
CC      -1- SIMILARITY: Belongs to the insulin family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way

```

```

CC modifiers and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; X61092; CAA3405.1; -.
CC DR PIR; B42179; B42179.
CC DR HSPB; P01308; IAI0.
CC DR InterPro; IPRO04825; Ins/IGF/relax.
CC DR Pfam; PF00049; Insulin; 1.
CC DR PRINTS; PRO0277; INSULINB.
CC DR SMART; SM00078; IGF; 1.
CC DR PROSITE; PS00262; INSULIN; 1.
CC KM Direct protein sequencing; Glucose metabolism; Hormone;
CC Insulin family; Signal.
CC FT SIGNAL 1 24
CC FT CHAIN 25 54 Insulin B chain.
CC FT PROPEP 57 87 C peptide.
CC FT CHAIN 90 110 Insulin A chain.
CC FT DISULFID 31 96 Interchain.
CC FT DISULFID 42 109 Interchain.
CC FT DISULFID 95 100
CC SQ SEQUENCE 110 AA; 12019 MW; 95A1F54BE7B247F9 CRC64;
CC
CC Query Match 98.5%; Score 456; DB 1; Length 110;
CC Best Local Similarity 98.8%; Pred.No. 3.9e-40;
CC Matches 85; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CC
CC Cy 1 FVNHQICGSHLYEALVYVCGRGFFYPKTRPREEDLVQVQVTLGGSGPAGSLQPLALEG 60
CC Db 25 FVNHQICGSHLYEALVYVCGRGFFYPKTRPREADPQVQVTLGGSPAGSLQPLALEG 84
CC
CC Cy 61 SLQKRGIVBQCTSIQSLVYQLENYCN 86
CC Db 85 SLQKRGIVBQCTSIQSLVYQLENYCN 110
CC
CC RESULT 9
CC INS_MACPA STANDARD; PRT; 110 AA.
CC ID INS_MACPA
CC AC P30406;
CC DT 21-JUL-1986 (Rel. 01, Created)
CC DT 13-AUG-1987 (Rel. 05, Last sequence update)
CC DT 05-JUN-2004 (Rel. 44, Last annotation update)
CC DE Insulin Precursor.
CC GN Name=INS;
CC OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
CC OC Cercopithecinae; Macaca.
CC OC NCBI_TaxID=9541;
CC RN [1]
CC RP MEDLINE=83080474; PubMed=6184262;
CC RA Metekam W., Gronenberg U., Leineweber M., Mengemayer F.,
CC RA Wilmacker E.-L.;
CC RT "The nucleotide sequence of cDNA coding for preproinsulin from the
CC RT primate Macaca fascicularis."
CC RL Gene 19:179-183(1982).
CC CC -1- FUNCTION: Insulin decreases blood glucose concentration. It
CC CC increases cell permeability to monosaccharides, amino acids and
CC CC fatty acids. It accelerates glycocolysis, the pentose phosphate
CC CC cycle, and glycogen synthesis in liver.
CC CC -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
CC CC disulfide bonds.
CC CC -1- SUBCELLULAR LOCATION: secreted.
CC CC -1- SIMILARITY: Belongs to the insulin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

or send an email to license@lsb-sib.ch.

CC EMBL: J00336; AAA35849.1; -

DR PIR: J00178; J00178.

DR HSPF: P01308; IAI0.

DR InterPro: IPR004825; Ins/IGF/relax.

DR Pfam: PF00049; Insulin.1.

DR PRINTS: PR00277; INSULINB.

DR SMART: SM00078; IIGF.1.

DR PROSITE: PS00262; INSULIN.1.

DR Glucose metabolism; Hormone; Insulin family; Signal.

KM SIGNAL

FT CHAIN 1 24

FT PROPEP 25 54 Insulin B chain.

FT CHAIN 57 87 C peptide.

FT CHAIN 90 110 Insulin A chain.

FT DISULFID 31 96 Interchain.

FT DISULFID 43 109 Interchain.

FT DISULFID 95 100 Interchain.

SO SEQUENCE 110 AA; 11991 MW; 83CEE33A80A420F9 CRC64;

Query Match 98.5%; Score 456; DB 1; Length 110;

Best Local Similarity 98.8%; Pred. No. 3.9e-40;

Matches 85; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVLYVCGERGFFYPTKRRRAEDLVQVYELGGPGAGSLQPLALEG 60

DB 25 FVNHLCGSHLVLYVCGERGFFYPTKRRRAEDLVQVYELGGPGAGSLQPLALEG 84

QY 61 SLOKRGIVEOCCSTICSLYQLENYCN 86

DB 85 SLOKRGIVEOCCSTICSLYQLENYCN 110

RESULT 10

INS_RABIT STANDARD; PRT; 110 AA.

AC P01311;

DT 21-JUL-1986 (Rel. 01, Created).

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Insulin precursor.

GN Name=INS;

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

CX NCBI_TaxID=9986;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=New Zealand white; TISSUE=Pancreas;

RX MEDLINE=94179230; PubMed=8132571;

RA Devaskar S.U., Giddings S.J., Rajakumar P.A., Carnaghi L.R.,

RA Menon R.K., Zaim D.S.;

RT "Insulin gene expression and insulin synthesis in mammalian neuronal

RT cells.";

RL J. Biol. Chem. 269:8445-8454(1994).

RN [2]

RP SEQUENCE OF 25-54 AND 90-110.

RX MEDLINE=66160119; PubMed=5949593;

RA Smith L.F.;

RT "Species variation in the amino acid sequence of insulin.";

RL Am. J. Med. 40:662-666(1966).

RN [3]

RP SEQUENCE OF 56-110 FROM N.A.

RC Giddings S.J., Carnaghi L.R., Devaskar S.U.;

RT Submitted (APR-1991) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: Insulin decreases blood glucose concentration. It

CC increases cell permeability to monosaccharides, amino acids and

CC fatty acids. It accelerates glycogenesis in liver.

CC cycle, and glycogen synthesis in liver.

CC -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two

CC disulfide bonds.

CC -1- SUBCELLULAR LOCATION: secreted.

CC -1- SIMILARITY: Belongs to the insulin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>

CC or send an email to license@lsb-sib.ch).

DR EMBL: U03610; AAA19033.1; -

DR EMBL: M61153; AAA17540.1; -

DR PIR: AS3438; INPB

DR HSPF: P01308; IGV6.

DR InterPro: IPR004825; Ins/IGF/relax.

DR Pfam: PF00049; Insulin.1.

DR PRINTS: PR00277; INSULINB.

DR SMART: SM00078; IIGF.1.

DR PROSITE: PS00262; INSULIN.1.

KM Direct protein sequencing; Glucose metabolism; Hormone;

INSulin family; Signal.

FT SIGNAL 1 24

FT CHAIN 25 54 Insulin B chain.

FT PROPEP 57 87 C peptide.

FT CHAIN 90 110 Insulin A chain.

FT DISULFID 31 96 Interchain.

FT DISULFID 43 109 Interchain.

FT DISULFID 95 100 Interchain.

FT CONFLICT 83 83 E -> Y (in Ref. 3).

SO SEQUENCE 110 AA; 11838 MW; 82D2975585D77FA8 CRC64;

Query Match 91.6%; Score 424; DB 1; Length 110;

Best Local Similarity 90.7%; Pred. No. 9.2e-37;

Matches 78; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVLYVCGERGFFYPTKRRRAEDLVQVYELGGPGAGSLQPLALEG 60

DB 25 FVNHLCGSHLVLYVCGERGFFYPTKRRRAEDLVQVYELGGPGAGSLQPLALEG 84

QY 61 SLOKRGIVEOCCSTICSLYQLENYCN 86

DB 85 SLOKRGIVEOCCSTICSLYQLENYCN 110

RESULT 11

INS_CANFA STANDARD; PRT; 110 AA.

AC P01321;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Insulin precursor.

GN Name=INS;

OS Canis familiaris (Dog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

CX NCBI_TaxID=9615;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=83109071; PubMed=6296142;

RA Kwok S.C.M., Chan S.J., Steiner D.F.;

RT "Cloning and nucleotide sequence analysis of the dog insulin gene.

RT Coded amino acid sequence of canine preproinsulin predicts an

RT additional C-peptide fragment.";

RL T. Biol. Chem. 258:2357-2363(1983).

RN [2]

RP SEQUENCE OF 25-54 AND 90-110.

RX MEDLINE=66160119; PubMed=5949593;

RA Smith L.F.;

RT "Species variation in the amino acid sequence of insulin.";

RL Am. J. Med. 40:662-666(1966).

CC -1- FUNCTION: Insulin decreases blood glucose concentration. It

CC increases cell permeability to monosaccharides, amino acids and

CC fatty acids. It accelerates glycogenesis, the pentose phosphate


```
CC cycle, and glycogen synthesis in liver.
CC -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
CC disulfide bonds.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the insulin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: V00179; CA23475.1; -.
CC PIR: A92413; IPDG.
CC HSSP: P01317; IARH.
CC InterPro: IPR004825; Ins/IGF/relax.
CC Pfam: PF00049; Insulin; 1.
CC PRINTS: PR00277; INSULINB.
CC SMART: SM00078; IIGF; 1.
CC PROSITE: PS00262; INSULIN; 1.
CC Direct protein sequencing; Glucose metabolism; Hormone;
CC Insulin family; Signal.
CC SIGNAL
CC FT CHAIN 1 24
CC FT PROPEP 57 54 Insulin B chain.
CC FT CHAIN 90 87 C peptide.
CC FT DISULFID 31 96 Insulin A chain.
CC FT DISULFID 43 109 Interchain.
CC FT DISULFID 95 100 Interchain.
CC SQ SEQUENCE 110 AA; 12190 MW; A5747918644FB98 CRC64;

Query Match 90.1%; Score 417; DB 1; Length 110;
Best Local Similarity 89.5%; Pred. No. 5e-36;
Matches 77; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYPTKTRREAEDLQVGGVGLGPGAGSLQPLALEG 60
DB 25 FVNQHLGSHLVEALYLVCGERGFFYPTKTRREAEDLQVGGVGLGPGAGSLQPLALEG 84
QY 61 SLQKRGIVGQCTSIQSLYOLENYCN 86
DB 85 ALQKRGIVGQCTSIQSLYOLENYCN 110

RESULT 12
INS_SPECTR STANDARD; PRT; 110 AA.
AC 091X13;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Insulin precursor.
GN Name=INS;
OS Spermophilus tridecemlineatus (Thirteen-lined ground squirrel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;
OC Spermophilus.
OC NCBI_TaxID=43179;
RN [1]
RP SEQUENCE FROM N.A.
RA Tiedra M.M., Buck M.J., Gubanyi J., Squire T.L., Andrews M.T.;
RT "Regulation of PDK4 expression in a hibernating mammal.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Insulin decreases blood glucose concentration. It
CC increases cell permeability to monosaccharides, amino acids and
CC fatty acids. It accelerates glycolysis, the pentose phosphate
CC cycle, and glycogen synthesis in liver.
CC -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
CC disulfide bonds.
CC -1- SUBCELLULAR LOCATION: Secreted.
```

```
CC -1- SIMILARITY: Belongs to the insulin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AY036604; AAK72558.1; -.
CC HSSP: P01308; IEV6.
CC InterPro: IPR004825; Ins/IGF/relax.
CC Pfam: PF00049; Insulin; 1.
CC PRINTS: PR00277; INSULINB.
CC SMART: SM00078; IIGF; 1.
CC PROSITE: PS00262; INSULIN; 1.
CC Glucose metabolism; Hormone; Insulin family; Signal.
CC SIGNAL
CC FT CHAIN 1 24
CC FT PROPEP 57 54 Insulin B chain.
CC FT CHAIN 90 87 C peptide.
CC FT CHAIN 31 96 Insulin A chain.
CC FT DISULFID 43 109 Interchain (By similarity).
CC FT DISULFID 95 100 Interchain (By similarity).
CC SQ SEQUENCE 110 AA; 12004 MW; 4511768D622BES5 CRC64;

Query Match 89.2%; Score 413; DB 1; Length 110;
Best Local Similarity 89.5%; Pred. No. 1.3e-35;
Matches 77; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYPTKTRREAEDLQVGGVGLGPGAGSLQPLALEG 60
DB 25 FVNQHLGSHLVEALYLVCGERGFFYPTKSRREVEEQGVGLGPGAGSLQPLALEM 84
QY 61 SLQKRGIVGQCTSIQSLYOLENYCN 86
DB 85 ALQKRGIVGQCTSIQSLYOLENYCN 110

RESULT 13
INS_HORSE STANDARD; PRT; 86 AA.
AC P01310;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Insulin precursor.
GN Name=INS;
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OC NCBI_TaxID=9796;
RN [1]
RP SEQUENCE OF 1-30 AND 66-86.
RA Harris J.I., Sanger F., Naughton M.A.;
RT "Species differences in insulin.";
RL Arch. Biochem. Biophys. 65:427-438 (1956).
CC -1- FUNCTION: Insulin decreases blood glucose concentration. It
CC increases cell permeability to monosaccharides, amino acids and
CC fatty acids. It accelerates glycolysis, the pentose phosphate
CC cycle, and glycogen synthesis in liver.
CC -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
CC disulfide bonds.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the insulin family.
```

-1- CAUTION: X's at positions 31-32 and 64-65 represent paired basic residues assumed by homology to be present in the precursor molecule.

DR PIR, A01580, IPR0.

DR HSBP, P01317, IAPB.

DR InterPro: IPR004825; Ins/IGF/relax.

DR Pfam: PF00049; Insulin, 1.

DR PRINTS: PR00277; INSULIN.

DR SMART: SMO0078; IIGF, 1.

DR PROSITE: PS00262; INSULIN, 1.

KW Direct protein sequencing; Glucose metabolism; Hormone; Insulin family.

FT CHAIN 1

FT PROPEP 33

FT CHAIN 66

FT DISULFID 72

FT DISULFID 19

FT DISULFID 71

FT DISULFID 76

SQ SEQUENCE 66 AA; 9142 MW; A3B1E822711BDA6 CRC64;

Query Match 85.1%; Score 394; DB 1; Length 66;

Best Local Similarity 84.9%; Pred. No. 1e-33; Mismatches 1; Indels 0; Gaps 0;

Matches 73; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 1 FVNHGCGSHLVATLVCGERGFYTPKTRREAEADLVGVGVEIGGFGAGSLQPLALEG 60

DB 1 FVNHGCGSHLVATLVCGERGFYTPKTRREAEADLVGVGVEIGGFGAGSLQPLALEG 60

QY 61 SLOKRGIVBOCCCTCSISLVLENYCN 86

DB 61 PQQKRGIVBOCCCTCSISLVLENYCN 86

RESULT 14

INS2_MOUSE STANDARD; PRT; 110 AA.

AC P01326;

DT 21-JUL-1986 (Rel. 01, Created)

DT 13-AUG-1987 (Rel. 05, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Insulin 2 precursor.

OS Name=Ins2; Synonyms=Ins-2;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=87169766; PubMed=3104603;

RA Wentworth B.M., Schaefer I.M., Villa-Komaroff L., Chirgwin J.M.;

RT "Characterization of the two nonallelic genes encoding mouse preproinsulin.";

RL J. Mol. Evol. 23:305-312(1986).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=NON;

RX MEDLINE=90372989; PubMed=2397023;

RA Sawa T., Ohgaki S., Morika H., Yano S.;

RT "Molecular cloning and DNA sequence analysis of preproinsulin genes in the NON mouse, an animal model of human non-obese, non-insulin-dependent diabetes mellitus.";

RL J. Mol. Endocrinol. 5:61-67(1990).

RN [3]

RP SEQUENCE OF 25-54 AND 90-110.

RX MEDLINE=72189455; PubMed=5063718;

RA Buenzli H.F., Glatthar B., Kunz P., Muehlaupt E., Humbel R.E.;

RT "Amino acid sequence of the two insulins from mouse (Mus musculus).";

RL Hoppe-Seyler's Z. Physiol. Chem. 353:451-456(1972).

CC -1- FUNCTION: Insulin decreases blood glucose concentration. It increases cell permeability to monosaccharides, amino acids and fatty acids. It accelerates glycolysis, the pentose phosphate cycle, and glycogen synthesis in liver.

CC -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two

CC disulfide bonds

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- SIMILARITY: Belongs to the insulin family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC

DR EMBL: X04724; CA26433.1; ..

DR PIR: A26342; INMS2.

DR HSBP, P01317, IAPB.

DR MGD; MGI:96573; Ins2.

DR GO; GO:0005634; C:nucleus, IDA.

DR GO; GO:0005732; C:small nuclear ribonucleoprotein complex, IDA.

DR GO; GO:0000187; P:activation of MAPK, IDA.

DR GO; GO:0006006; P:glucose metabolism, IMP.

DR GO; GO:0008285; P:insulin receptor signaling pathway, IDA.

DR GO; GO:0016042; P:lipid catabolism, IDA.

DR GO; GO:0042981; P:regulation of apoptosis, IMP.

DR GO; GO:0042325; P:regulation of phosphorylation, IDA.

DR GO; GO:0006983; P:response to ER overload, IMP.

DR InterPro: IPR004825; Ins/IGF/relax.

DR Pfam: PF00049; Insulin, 1.

DR PRINTS: PR00277; INSULIN.

DR SMART: SMO0078; IIGF, 1.

DR PROSITE: PS00262; INSULIN, 1.

KW Direct protein sequencing; Glucose metabolism; Hormone; Insulin family; Multigene family; Signal.

FT SIGNAL 1

FT CHAIN 25

FT PROPEP 57

FT CHAIN 90

FT DISULFID 31

FT DISULFID 43

FT DISULFID 95

SQ SEQUENCE 110 AA; 12364 MW; 3554C880D24FAD6 CRC64;

Query Match 85.1%; Score 394; DB 1; Length 110;

Best Local Similarity 84.9%; Pred. No. 1.3e-33; Mismatches 9; Indels 0; Gaps 0;

Matches 73; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 FVNHGCGSHLVATLVCGERGFYTPKTRREAEADLVGVGVEIGGFGAGSLQPLALEG 60

DB 25 FVNHGCGSHLVATLVCGERGFYTPKTRREAEADLVGVGVEIGGFGAGSLQPLALEG 60

QY 61 SLOKRGIVBOCCCTCSISLVLENYCN 86

DB 61 SLOKRGIVBOCCCTCSISLVLENYCN 86

QY 85 AQQKRGIVBOCCCTCSISLVLENYCN 110

DB 85 AQQKRGIVBOCCCTCSISLVLENYCN 110

RESULT 15

INS2_RAT STANDARD; PRT; 110 AA.

AC P01323;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Insulin 2 precursor.

OS Name=Ins2; Synonyms=Ins-2;

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=Sprague-Dawley; TISSUE=liver;

RX MEDLINE=80045035; PubMed=4962284;

RA Lomedico P., Rosenthal N., Efstratiadis A., Gilbert W., Kolodner R.,

RT "The structure and evolution of the two nonallelic rat preproinsulin
 RT genes.";
 RL Cell 18:545-558(1979).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86310682; PubMed=2427930;
 RA Soares M.B., Schin E., Henderson A., Karathanasis S.K., Cate R.,
 Zeilin S., Chirgwin J., Efstratiadis A.;
 RT "RNA-mediated gene duplication: the rat preproinsulin I gene is a
 functional retroposon.";
 RL Mol. Cell. Biol. 5:2090-2103(1985).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80240379; PubMed=6249167;
 RA Lomedico P.T., Rosenthal N., Kolodner R., Efstratiadis A., Gilbert W.;
 RT "The structure of rat preproinsulin genes.";
 RL Ann. N. Y. Acad. Sci. 343:425-432(1980).
 RN [4]
 RP SEQUENCE OF 25-54 AND 90-110.
 RX MEDLINE=70067613; PubMed=4311938;
 RA Steiner D.F., Clark J.L., Nolan C., Rubenstein A.H., Margolis E.,
 Aten B., Oyer P.E.;
 RT "Proinsulin and the biosynthesis of insulin.";
 RL Recent Prog. Horm. Res. 25:207-282(1969).
 RN [5]
 RP SEQUENCE OF 57-87.
 RX MEDLINE=73061498; PubMed=4640931;
 RA Tager H.S., Steiner D.F.;
 RT "Primary structures of the proinsulin connecting peptides of the rat
 and the horse.";
 RL J. Biol. Chem. 247:7936-7940(1972).
 RN [6]
 RP SEQUENCE OF 57-87, AND REVISIONS.
 RX MEDLINE=7217385; PubMed=4554104;
 RA Markussen J., Sundby F.;
 RT "Rat-proinsulin C-peptides. Amino-acid sequences.";
 RL Eur. J. Biochem. 25:153-162(1972).
 CC -I- FUNCTION: Insulin decreases blood glucose concentration. It
 increases cell permeability to monosaccharides, amino acids and
 fatty acids. It accelerates glycolysis, the pentose phosphate
 cycle, and glycogen synthesis in liver.
 CC -I- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
 disulfide bonds.
 CC -I- SUBCELLULAR LOCATION: Secreted.
 CC -I- SIMILARITY: Belongs to the insulin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 or send an email to license@isb-sdb.ch).
 CC -----
 CC EMBL: V01243; CAA24560.1; -;
 DR EMBL: J00748; AAA41443.1; -;
 DR EMBL: M25585; AAA41440.1; -;
 DR EMBL: M25583; AAA41440.1; JOINED.
 DR PIR: B80789; IPR2.
 DR HSRF: P01317; IAPH.
 DR RGD: 2916; Ins2.
 DR InterPro: IPR004825; Ins/IGF/relax.
 DR Pfam: PF00049; Insulin; 1.
 DR PRINTS: PRO0277; INSULINB.
 DR SMART: SM00078; IIGF; 1.
 DR PROSITE: PS00262; INSULIN; 1.
 KW Direct protein sequencing; Glucose metabolism; Hormone;
 KW Insulin family; Multigene family; signal.
 ZI SIGNAL 1 24
 FT CHAIN 25 54 Insulin 2 B chain.
 FT PROPEP 57 87 C peptide.
 FT CHAIN 90 110 Insulin 2 A chain.
 FT DISULFID 31 96 Interchain.

FT DISULFID 43 109 Interchain.
 FT DISULFID 95 100
 SO SEQUENCE 110 AA; 12339 MW; 3A626DA98C86F3CA CRC64;
 Query Match 85.1%; Score 384; DB 1; Length 110;
 Best Local Similarity 84.9%; Pred. No. 1.3e-33; Indels 0; Gaps 0;
 Matches 73; Conservative 4; Mismatches 9;
 Db 25 FVQKHLGGSHLVEALYLVGGERGFFYTPKTREAEDELQVGVELGGGPGAGSLQPLALEG 60
 FVQKHLGGSHLVEALYLVGGERGFFYTPKTSRREVEDPQVALTELGCGGPGAGDLQTLALEV 84
 QY 61 SLQKRGIVQCCTSIQSLYOLENYCN 86
 Db 85 ARQKRGIVQCCTSIQSLYOLENYCN 110

Search completed: November 2, 2004, 20:20:33
 Job time : 85.3026 secs

This Page Blank (uspto)